

Exploring the Evidence for Email Phishing Training: A Scoping Review



*This report is submitted in partial fulfilment of the degree of Master of Psychology
(Organisational and Human Factors)*

School of Psychology

University of Adelaide

October 2023

Word Count: 7,990

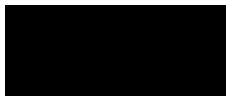
Table of Contents

Declaration	iii
Statement of Contribution	iv
Acknowledgements	v
Journal Formatting Requirements	vi
Abstract	vii
Method	5
Protocol and Registration	6
Literature Sources and Search Strategy (Stage 2)	6
Study Criteria, Screening, and Selection (Stage 3)	7
Data Charting, Extraction and Synthesis (Stages 4 and 5)	9
Results	9
Characteristics of Included Studies	9
Anti-Phishing Training Delivery	10
Embedded Training	19
Standalone Training via Passive Modalities	22
Standalone Training via Interactive Modalities	23
Anti-Phishing Training Content and Pedagogical Approach	28
Rules-Based Training	29
Framing Effects as Persuasive Appeals to Enhance Training Outcomes	29
Experiential Learning	29
Mindful Awareness Training	30
Discussion	30
General Summary of Findings	31
Training Modalities Empirically Evaluated	33
Features of Training Delivery Associated with Improved Outcomes	34
The Influence of Training Content and Pedagogical Approach	36
Limitations and Future Directions	37
Conclusion	38
References	40
Appendix A	55
Appendix B	56

Declaration

This dissertation contains no material which has been accepted for the award of any other degree or diploma in any University, and, to the best of my knowledge, contains no materials previously published except where due reference is made. I give permission for the digital version of my dissertation to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the School to restrict access for a period of time.

Signature:



Student ID:



Month/Year: October 2023

Statement of Contribution

In writing this thesis, my supervisors and I used a collaborative approach to formulate the research questions and the study design, including the selection of an appropriate methodology. I held sole accountability for the development of the study eligibility and screening criteria, collection and analysis of data, interpretation of the findings, and preparation/writing of the report. A library research assistant reviewed the logic grids for completeness, and supervisor 1 reviewed a sub-sample of abstracts to independently confirm reliability of study screening. Lastly, supervisors 1 and 2 jointly supervised me throughout this process and provided review and editing of drafts.

Acknowledgements

I am extremely grateful to my supervisors, Dr. xxx and Dr. xxx for their generosity in sharing their time and immense knowledge with me over the course of this work. Your invaluable support, guidance and feedback has certainly pushed me to be better at what I do. I feel privileged to have had the opportunity to learn a great deal from you both and wish to thank you for making this experience an enjoyable one.

Journal Formatting Requirements

Computers & Security has been selected as the notional journal for publication, being one of the most respected journals for IT security research globally. There are no strict formatting requirements required for submissions, although all manuscripts must be divided into clearly defined sections and contain the essential elements of a research report (i.e., Abstract, Keywords, Introduction, Materials and Methods, Results, and Conclusions; see Appendix A). Similarly, there are no strict requirements on reference formatting, which may be in any style or format as long as the style is consistent. Therefore, I have formatted to APA7 standards.

Abstract

Background: Phishing emails are a pervasive threat to the security of confidential information worldwide. To mitigate this risk, a wide range of training measures have been developed to target the human factors involved in phishing email susceptibility. Despite the importance and widespread use of anti-phishing training programs, there is no clear understanding of the various approaches that are used, and the extent to which these approaches have been assessed. *Objective:* The primary aim of this scoping review was to identify and describe the nature of available training interventions and their measurable outcomes on user susceptibility, as reported in published articles. *Methods:* Systematic searches using predefined keywords within PsycINFO, PubMed (MEDLINE) and Web of Science identified 42 studies that met the inclusion criteria. Each included study was critically analysed, and a standardised data extraction spreadsheet used to systemise the data that informed the descriptive narrative review. *Results:* Findings revealed that near-term training impact is well documented, however evidence on the success of programs in driving sustained behavioural change is limited. Components of training design influencing the effectiveness of outcomes included training intensity, active approaches to learning, the provision of detailed feedback, and supplementing attentional awareness skills-based training with traditional cue-based approaches. *Conclusions:* Improved user resilience to phishing emails confirms the utility of training as an important defensive mechanism, although current approaches leave approximately 20% of users at risk. Findings provide useful clarity in respect of what is known and where there are prominent gaps in the evidence base, alongside directions for future research.

Keywords: phishing email, phishing susceptibility, training, cybersecurity, human cognition

Exploring the Evidence for Email Phishing Training: A Scoping Review

With reliance on digital communication technologies now indispensable to business activities, the integrity of a large volume of information is vulnerable to exploitation by cybercriminals. Phishing represents one of the most pervasive forms of cybercrime, involving the use of deceptive techniques to convince users to divulge sensitive information (Alkhalil et al., 2021). Typically, attackers exploit people's familiarity and trust in major brands or institutions to lure users to click on malicious links or attachments embedded within emails. Cyberattacks may occur through various mediums (e.g., SMS/text phishing), however the most common means is via phishing emails (ProofPoint, 2023). Growing at a rate of more than 150% per year since 2019, email-based attacks are becoming increasingly common, reaching unprecedented levels with 4.7 million attacks logged in 2022 alone (Anti Phishing Working Group, 2022). In parallel, the cost of investigating and remediating a data breach is becoming increasingly expensive (IBM Security, 2023), involving direct financial losses alongside prolific costs associated with operational disruption, reputational damage, legal and regulatory penalties (Anderson et al., 2013).

Despite advancements in technological countermeasures such as email filters and blacklists (Aleroud & Zhou, 2017), these methods cannot deliver a complete solution (Furnell & Clarke, 2012). The increasing sophistication of phishing attacks (Gupta et al., 2017) means many emails continue to evade automated systems, with a substantive 85% of security breaches attributable to the vulnerabilities of user error (ProofPoint, 2022; Vayansky & Kumar, 2018). It is therefore widely recognised that a comprehensive defence strategy must consider how best to reduce the human side of cybersecurity risk when automation fails (Furnell & Clarke, 2012; Heartfield et al., 2016).

User training has emerged as an integral component of an organisation's overall security posture, aiming to equip individuals with the skills to recognise and respond to phishing

attempts (Desolda, 2022). While training can reduce the success rate of attacks (Caldwell, 2016; Kumaraguru et al., 2010), many individuals continue to fall victim to phishing emails after having received training (Caputo et al., 2014; Wash & Cooper, 2018). For example, Sheng et al. (2010) found that even though the rate of phishing victimisation reduced by 40% post training, participants still failed to detect phishing emails 28% of the time. Additionally, a recent study revealed that of the organisations who fell victim to phishing attacks, 65% had previously trained their staff (Cloudian, 2021). The utility of training as an effective mitigant against phishing therefore remains a subject of debate (Caldwell, 2016; Khonji et al., 2013). While methodological inconsistencies may certainly contribute to variable findings (e.g., sampling error, operationalisation of variables), mixed results are more likely to derive from the fact that approaches to phishing training are greatly varied.

One common approach to training is knowledge-based instruction. From a theoretical perspective, knowledge training should enhance performance in novel situations where prior rules or experience are not available for guidance (Rasmussen, 1986). However, knowledge-based performance is both slow and cognitively demanding thus atypically favoured within real world contexts (e.g., under conditions of time constraints and uncertainty; Klein, 2008). Moreover, knowledge alone is insufficient for the development of expertise within a given domain (Kahneman & Klein, 2009). In contrast, *skilled behaviour* – characterised by automaticity and effortless execution - requires the accumulation of knowledge alongside repeated, deliberate practice and experience (Ericsson 2008; Rasmussen, 1986) such that learned knowledge can be retained, recalled, and skilfully applied in practical settings (Burke & Hutchins, 2007; Grossman & Salas, 2011; Salas & Cannon-Bowers, 2001).

It has been suggested that one of the reasons phishing training may fail to yield intended results is because the methods employed often emphasise knowledge enhancement, expecting behaviour change to naturally follow (Albrechtsen & Hovden, 2010; Jaeger et al., 2021; van

Steen et al., 2020). Currently there is no holistic review of the evidence that has sought to clarify whether this is indeed the case. While several studies have reported a positive association between phishing knowledge and phishing resilience (Downs et al., 2006; Jakobbsen et al., 2007; Zafar et al., 2019), in some cases, greater phishing knowledge has been associated with an *increase* in susceptibility (Anandpara et al., 2007; Diaz et al., 2020). A lack of clarity around the extent to which training has reduced user susceptibility to phishing emails therefore warrants investigation.

Another hurdle training programs need to overcome is the consideration of the various cognitive biases that may guide decisions (Tversky & Kahneman, 1974). Phishing typically succeeds when attackers are able to exploit cognitive biases through social engineering techniques such as persuasive appeals to authority or urgency (Downs et al, 2006; Parsons et al., 2019; Zielinska et al., 2014). Such appeals may, for example, convince users to respond without deliberation, preferencing rapid heuristic thinking over more rigorous systematic processing (Tversky & Kahneman, 1974). Results generally support the notion that users fall victim to phishing emails due to a reliance on a limited set of superficial cues rather than a broader range of content features within emails, making it less likely that deceptive cues are detected (Vishwanath et al., 2011; Vishwanath et al., 2016). The importance of cues in guiding expert decision making is well established in a number of operational environments, including firefighting and medical diagnosis (Klein, 1986; Klein & Calderwood 1991). Within these domains, targeted cue-based training, through repeated engagement with case examples, has proven efficacy in the enhancement of task performance (Kahneman & Klein, 2009). However, the level of expertise identified in the examples above result from *years* of exposure. Given that anti-phishing training is usually conducted over a limited number of days or hours, it is unclear whether such training is adequate for the development of meaningful cues. Nonetheless, various researchers have advocated the use of cue inventories

as the basis of phishing training interventions (Sturman et al., 2023; Wiggins & O'Hare, 2003), yet it is unknown whether this method is in fact used or has proven effectiveness when it comes to phishing.

Training impact may also be moderated by the choice of training design and implementation. Numerous anti-phishing training programs have been commercialised which follow different pedagogical approaches, from traditional instructor-led training to more recent gamified approaches (Jampen et al., 2020) as well as variations in training content, from succinct rules-based instructions (Kumaraguru et al., 2010) to more elaborated, story-based forms (e.g., Wash & Cooper, 2018). Evidence within similar domains suggests that the choice of training design and implementation method influences training utility. For example, a recent meta-analysis found that an inconsistent approach to the development of security education training and awareness programs (SETA) partially explains variable findings related to their efficacy (Cram et al., 2019). Similarly, a literature review of SETA programs found that hands-on delivery methods contributed to training success, while excessive information and extensive use of multi-media were counterproductive (Hu et al., 2022). Comparable reviews are required in the context of email phishing; currently, there is no clear evidence of the breadth of the various training approaches that are used nor the extent to which these approaches have been assessed. Consequently, there is limited understanding of how best to design anti-phishing training programs to maximise their utility. Despite this lack of understanding, organisations are placing great reliance and investment in such programs to manage their cybersecurity risk (Telstra Corporation, 2018).

To address these questions, a comprehensive summary of the empirical evidence regarding the nature and outcomes of training programs is required. Prior reviews have addressed some of these aspects, but not all. Specifically, Jampen et al. (2020) provide a valuable summary of key findings from numerous anti-phishing training studies, however,

they did not critically review the impact of training design on user-susceptibility to identify common indicators of success, nor did they consider training content as a construct of interest. Similarly, Abawajy (2014) examined the effectiveness of various delivery approaches on security awareness outcomes, although confined their research efforts to game-, text-, and video-based methods, thus omitting the most frequently used methods today, being e-learning, simulation-based exercises, and instructor-led programs (ProofPoint, 2023). Our study diverges from past reviews in that we aim to provide a more comprehensive and current perspective on the collective body of work in this field, which can provide practical contributions to both research and practice alike for the development of optimal training solutions.

Against this background, the aim of this scoping review is to identify and describe:

- i) the nature of training programs that have been empirically assessed (i.e., what content is being trained and how is this content being delivered?)
- ii) the evidence for the efficacy of the training programs identified (i.e., what outcome variables have been assessed and what can be reasonably concluded?),
and
- iii) the aspects of training design associated with improved outcomes.

We hope to provide useful clarity about what is currently known from the available literature, and what remains to be further explored in support of the design of targeted and efficacious training solutions.

Method

We aligned our methods to the first five stages of undertaking a scoping review as identified by Levac et al. (2010) and reported in accordance with the Cochrane and Preferred Reporting Items for Systematic Review and Meta-Analyses Extension for Scoping Reviews guidelines (PRISMA-ScR; Tricco et al., 2018; see Appendix B). After assembling the

research team, we confirmed the purpose of the study and associated research questions to guide the scope of inquiry (Stage 1).

Protocol and Registration

The study protocol has been registered in the Open Science Framework (registration: [removed for blind review]).

Literature Sources and Search Strategy (Stage 2)

Articles published in peer reviewed journals from 1 January 2003 to 31 March 2023 were sourced from three databases: PsycINFO, PubMed (MEDLINE) and Web of Science.

Eligible studies examining training interventions aimed at reducing end-user susceptibility to phishing emails were identified using a comprehensive list of key words and index terms developed in consultation with a senior research librarian to ensure broad coverage of available literature. The search strategy was designed in PsycINFO and translated to other databases (see Table 1). In addition, a hand-search of the reference lists of included studies and relevant narrative reviews (Alhashmi et al., 2021; Hu et al., 2022; Jampen et al., 2020) was undertaken, along with citation searching in Scopus.

The Anti-Phishing Work Group (APWG) was formed in 2003 to raise phishing awareness and develop effective countermeasures to thwart the rising occurrence of phishing attacks (APWG, 2022). Therefore, the time period from 2003 to 2023 was considered appropriate for our review.

Table 1

Search Strategy as Applied in PsycINFO

Term	Search String
1. Phishing	(Phish* OR spearphish* OR spear-phish* OR antiphish* OR anti-phish*).mp OR cybercrime.sh OR cybersecurity.mp OR "cyber security".mp AND
2. Training	(Training OR education OR awareness).mp

Study Criteria, Screening, and Selection (Stage 3)

To ensure feasibility and meet the practicalities of time constraints, only those articles published in English and in peer reviewed journals were included. Studies also had to meet the eligibility criteria outlined in Table 2, conceptualised using the Population Intervention Comparison Outcome Study design (PICO-S) framework (Robinson et al., 2013).

Table 2

Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Time Period	2003 – 2023	Articles outside this range
Language	English	Non-English studies
Population	Adult participants aged > 17years, whether students or those in an employment setting	Any papers whose participants included those aged < 17 years
Intervention	All programs whose primary aim was the reduction of user susceptibility to phishing emails, irrespective of content, duration or setting	Programs focused on reducing user susceptibility to non-email-based forms of phishing attacks (e.g., SMS), or those examining technical solutions.
Comparator	Any comparator was eligible (e.g., no training control, active control, pre/post studies)	Studies with no comparator condition
Outcome	Any objective or subjective measure of post-intervention variation in phishing susceptibility	Any other outcomes (e.g., usability preferences)
Study Type	Experimental, quasi-experimental, and observational user studies (e.g., randomized controlled trials, before-and-after studies, prospective or retrospective cohort studies). Peer-reviewed conference abstracts provided sufficient information is provided.	Non-empirical studies (e.g., conceptual narrative reviews, editorials, opinion pieces), or those reporting insufficient data to conduct analysis

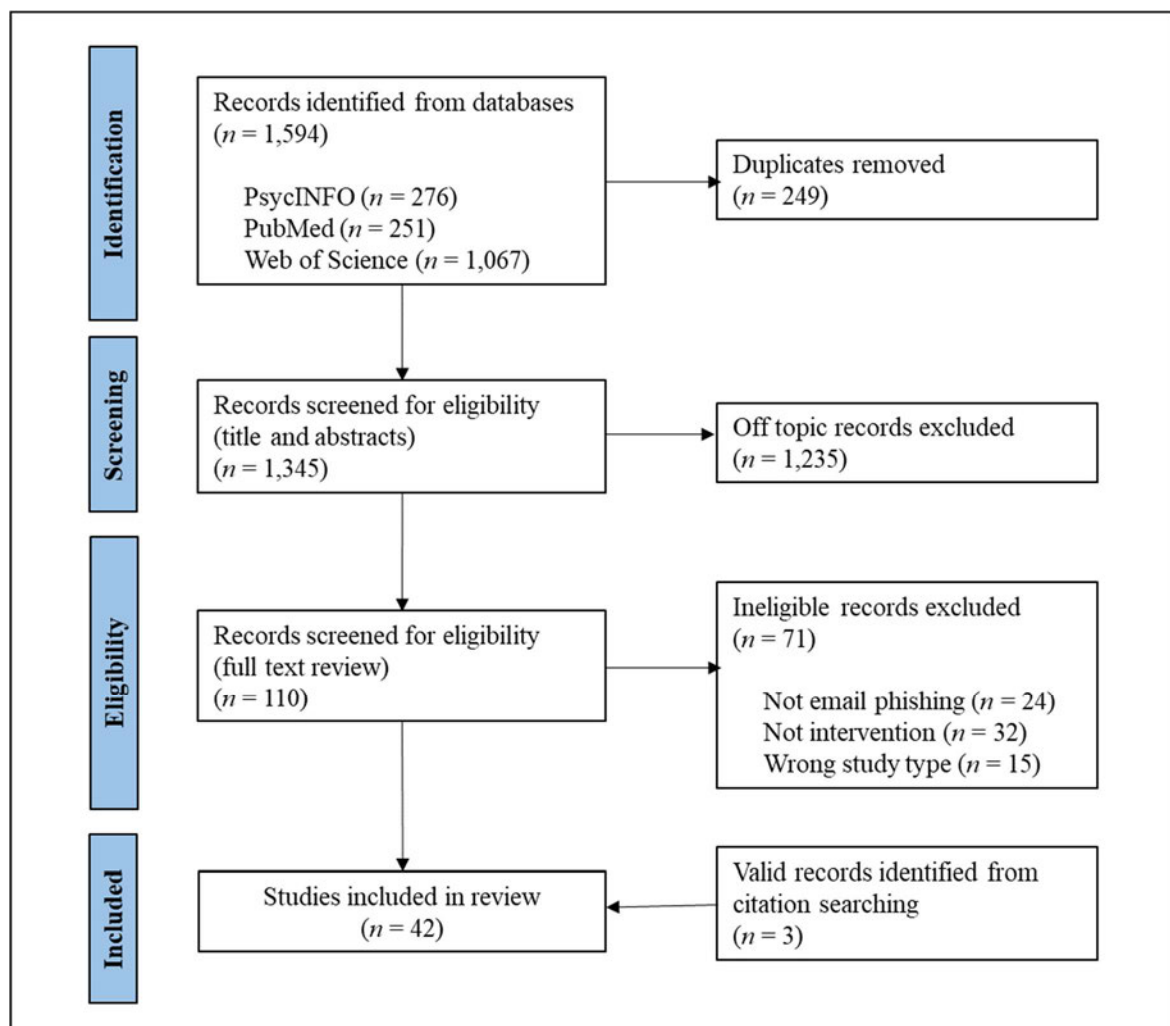
Screening was conducted using Covidence systematic review software (Veritas Health Innovation). A total of 1,594 studies were identified from the search process, with 249 duplicates automatically filtered by Covidence (see Figure 1). Each title and abstract

($n = 1,345$) were screened, which identified a large number of articles as irrelevant ($n = 1,235$). These articles were primarily associated with technological solutions or involved evaluations of user susceptibility, so were excluded.

The remaining 110 articles were screened by full text review, with 42 independent studies (k) identified as meeting all eligibility criteria. To ensure reliability of the screening process, a random subsample of full-text records ($n = 11$, 10%) were independently examined by a human factors researcher (Supervisor 1). Substantial agreement between raters was achieved ($K = 0.70$; Viera & Garrett, 2005) with discrepancies resolved through consensus discussion involving a third researcher (Supervisor 2).

Figure 1.

Flow Diagram Outlining Study Selection Process (adapted from PRISMA; Page et al., 2021)



Data Charting, Extraction and Synthesis (Stages 4 and 5)

In accordance with the PRISMA-ScR guidelines (Tricco et al., 2018), key data were retrieved from each study using a pre-piloted Microsoft Excel spreadsheet which was iteratively re-evaluated by the research team during the course data extraction to ensure it continued to meet the objectives of the scoping review. Extracted data included: 1) study characteristics (author, publication year, sample size, country of origin, study design, context); 2) sample demographics (mean age, age range, gender, recruitment source); 3) intervention characteristics (aim, training format, delivery mode, program content); 4) outcomes (source of measure, reported results, and training retention, i.e., length of time between the end of training and the outcome measured). During the course of data extraction it became clear that ‘study limitations’ was a variable of interest, with several studies indicating methodological limitations and areas in need of further research. This item was therefore added to the data extraction spreadsheet *post-hoc*. Data was extracted by the student and cross-referenced by a human factors researcher for consistency and accuracy.

As we were interested in describing the various types of training modalities, content, and outcomes documented in the available literature, we coded these variables against themed categories to investigate the occurrence of concepts. Using an inductive process, categories were derived from emergent patterns in the extracted data itself rather than from preconceived classifications. The resultant categories were then reviewed and synthesised into higher order themes, allowing the descriptive narrative review of results to be supplemented by frequency counts (see Tables 3 and 4).

Results

Characteristics of Included Studies

Our sample comprised 42 studies published between 2007 and 2023 (see Table 3). Typically, intervention outcomes were compared against no training controls ($k = 23$; 55%)

or were measured using pre/post designs ($k = 14$; 33%). The number of participants varied considerably between studies, ranging from 20 to in excess of 30,000 ($M = 3,001$, $SD = 6,977$), with smaller samples generally confined to lab-based experiments, and larger cohorts represented in field studies. Lab-based designs formed 50% of the study pool ($k = 21$) whose participants were predominantly volunteer undergraduate students ($k = 16$; 76%), while those from field studies ($k = 21$) included university staff and students ($k = 10$; 48%) or employees from various organisations across both public and private sectors ($k = 11$; 52%). The age range of participants across the dataset spanned 17 to 73 years, with a reasonably balanced representation between males and females for those that include gender as a demographic variable of interest ($k = 28$; 67%). More than half of the studies originated from the United States of America ($k = 26$; 62%).

Anti-Phishing Training Delivery

Training delivery methods are typically classified into two broad categories, namely persistent (or embedded training) and standalone methods. Standalone modalities can be further defined by the degree of user engagement (i.e., passive or interactive). Passive methods focus on developing user knowledge and understanding (i.e., written training materials and educational videos). In contrast, interactive methods typically involve hands-on practice where users can develop specific skills and receive real-time feedback (i.e., e-learning, educational games, and instructor-led training; see Table 4).

Table 3*Characteristics of Included Studies*

Study Name (Year)	N	Participants	Mean Age (Years)	Country	Training Condition(s)	Comparator(s)	Training Effect Outcome Measures	Context	Time of Assessment
Abawajy (2014)	60	Volunteers	-	Australia	Text verse Game verse Online video	Randomised Between Groups	User preferences; URL & website classification	Lab	Immediate Post Intervention
Arachchilage et al. (2016)	20	Undergraduate student volunteers	-	UK	Gamified	Pre/Post	User satisfaction; phishing website detection; self- assessed behaviours	Lab	Immediate Post Intervention
Back & Guerette (2021)	2000	University staff	-	USA	e-learning	No training control	Click through rates: open/click/submit	Field	> 4 weeks, ≤28 weeks
Burns et al. (2019)	260	MBA students	28	USA	Embedded (online written material; various framing cond.)	No training control	Click through rates	Field	> 4 weeks, ≤ 12 weeks
Caputo et al. (2014)	1359	Employees; US organisation	-	USA	Embedded (online written material)	No training control	Click through rates; time spent viewing training page; qualitative data	Field	> 4 weeks, ≤12 weeks
Carella et al. (2017)	150	Undergraduate student volunteers	-	USA	Embedded (online written material vs. in- person training)	No training control	Click through rates	Field	> 4 weeks, ≤12 weeks
Cuchta et al. (2019)	4,777	University students, faculty, staff	-	USA	Embedded (long doc vs. succinct visual doc vs. game)	No training control	Click through rates	Field	> 4 weeks, ≤ 12 weeks
Daengsi et al. (2022)	19,938	Employees; financial services	-	Thailand	Embedded (e-Learning + in-person workshop)	Pre/Post	Click through rates	Field	
Davinson & Sillence (2010)	64	Undergraduate student volunteers	22	UK	Gamification	No training control	Self-assessed behaviours	Lab	≤ 1 week

Study Name (Year)	N	Participants	Mean Age (Years)	Country	Training Condition(s)	Comparator(s)	Training Effect Outcome Measures	Context	Time of Assessment
Dodge et al. (2012)	892	Military personnel	-	USA	Embedded (notification only vs. notification + online training)	No training control	Click through rates	Field	> 4 weeks, ≤12 weeks
Gokul et al. (2018)	8,071	Employees; tech savvy	-	India	Gamification	Pre/Post	Correctness & confidence URL identification; user experience	Field	Immediate Post Intervention
Gordon et al. (2019)	5,416	Employees; healthcare	-	USA	e-Learning	Pre/Post	Click through rates	Field	> 26 weeks
Jansson & von Solms (2013)	25,579	University employees	-	South Africa	Embedded (warning notification + optional e-Learning)	Pre/Post	% disclosures; % completing optional training	Field	≤ 1 week
Jensen et al. (2017)	355	University students, faculty & staff	-	USA	Rule-based vs. mindful processing	No training control	% responses to simulated phishing email	Field	> 1 week, ≤ 4 weeks
Kavrestad et al. (2022)	41	University students & employees	-	Sweden	Game verse embedded prompt within email inbox	No training control	% correct email classifications; eye gaze behaviour	Lab	Immediate Post Intervention
Kim et al. (2020)	1248	Employees; governmental organisation	-	Korea	Lecture	No training control; punishment control	% phished	Field	> 4 weeks, ≤12 weeks
Kumaraguru et al. (2007a)	30	Volunteers (novice users)	21 (comic) 27 (text) 31 (notice)	USA	Embedded training page (various formats)	Between groups	Click through rates; user preferences; confidence indicators	Lab	Immediate Post Intervention
Kumaraguru et al. (2007b)	42	Volunteers recruited via flyers	25 (emb.) 24 (non.) 28 (contr.)	USA	Embedded training page vs. non-embedded training page	No training control	Mean number of correct responses for email classifications	Lab	≤ 1 week

Study Name (Year)	N	Participants	Mean Age (Years)	Country	Training Condition(s)	Comparator(s)	Training Effect Outcome Measures	Context	Time of Assessment
Kumaraguru et al. (2008)	301	Employees; organisation	-	Portugal	Embedded training page generic vs. embedded spear-phish training	No training control	Click through rates for those who disclosed sensitive information	Field	≤ 1 week
Lain et al. (2021)	14,773	Employees; private organisation	-	Switzerland	Embedded warning notification vs. training page	No warning notification control / No training control	Click through rates; submitted personal data; reported phishing emails	Field	
Lim et al. (2021)	42	Volunteers recruited via online survey	20	USA	Written material only vs. written material + e-learning	No training control	% correct email classification task; confidence scores	Lab	Immediate Post Intervention
Mayhorn & Nyeste (2012)	84	Volunteer Psychology students	19	USA	Embedded training page vs. game + embedded training page	No training control	Hits/misses/false alarms/correct rejections	Lab	≤ 1 week
McElwee et al. (2018)	1,000	Employees & contractors	-	USA	Institutional training (undefined)	Pre/post	Click through rates	Field	
Nguyen et al. (2021)	453	Undergraduate students	20	USA	Rule based training vs. mindful processing	No training control	Hits/misses/false alarms/correct rejections	Lab	> 4 weeks, ≤ 12 weeks
Reinheimer et al (2020)	409	Employees; public administration	-	Germany	In person tutorial with/without refresher training.	Pre/post	Hits/misses/false alarms/correct rejections	Field	> 26 weeks
Roepke et al. (2022)	89	Volunteer University students	-	Germany	Gamification	Pre/post	URL classification performance; confidence ratings; self-assessed behaviours	Lab	> 4 weeks, ≤ 12 weeks
Sarno et al. (2022)	75	Volunteer university students	19	USA	Classification aid (hardcopy) verse feedback	No training control	Hits/misses/false alarms/correct rejections	Lab	Immediate Post Intervention

Study Name (Year)	N	Participants	Mean Age (Years)	Country	Training Condition(s)	Comparator(s)	Training Effect Outcome Measures	Context	Time of Assessment
Sharevski & Jachim (2022)	120	Volunteers university students	-	USA	Voice assistant vs. facts and advice training	No training control	% correct for each persuasion principle	Lab	Immediate Post Intervention
Sheng et al. (2010)	1,001	Volunteers recruited via Amazon mTurk	30	USA	Text vs. Game vs. Comic Training Page vs. Game + Comic Training Page	No training control	% correct classification task (i.e., responded they would submit information)	Lab	Immediate Post Intervention
Silic & Lowry (2020)	420	Volunteers from international organisation	33	UK, USA, Australia	Gamification vs. online written material	No training control	% phished (opened email)	Field	
Singh et al. (2023)	296 Ex.1 224 Ex.2	Volunteers recruited via Amazon mTurk	35 (Exp 1) 37 (Exp 2)	USA	Feedback under low or high phishing email exposure	Minimal Feedback; No feedback	Hits/misses/false alarms/correct rejections; confidence scores; self-reported actions	Lab	Immediate Post Intervention
Stockhardt et al. (2016)	81	Volunteer students	17 (instr.) 20 (e-lrn) 22 (text)	Germany	Workshop vs. e-Learning vs. written material	Pre/Post & Between Groups	% correct email classifications; knowledge test; confidence; user experience	Lab	Immediate Post Intervention
Sumner et al. (2022)	110 pre 32 post	Volunteer University employees	-	USA	e-Learning	Pre/post	% accurate classification task classification; confidence; self-assessed knowledge / behaviour	Lab	Immediate Post Intervention
Sutter et al. (2022)	31,940	University students (81%) and staff (19%)	-	Switzerland	e-learning vs. written material vs. e-learning + written material	No training control	Click through rates (clicked / submitted); % completed training	Field	
Tschakert & Ngamsuriyaraj (2019)	34	Volunteer university students	20-21 median	Thailand	Instructor + self-paced training (text, video, game) vs. self-paced content only	Pre/post and Between groups	Hits/misses/false alarms/correct rejections; % correct classifications; confidence; satisfaction	Lab	≤ 1 week

Study Name (Year)	N	Participants	Mean Age (Years)	Country	Training Condition(s)	Comparator(s)	Training Effect Outcome Measures	Context	Time of Assessment
Volkamer et al. (2018)	89	Volunteers recruited online	36 pre/post 38 delay	Germany	Video	Pre/post	% correct email discrimination task	Lab	> 4 weeks, ≤ 12 weeks
Wash & Cooper (2018)	1,945	University staff	-	USA	Text (stories vs. facts & advice) x delivery (expert vs. non-expert)	No training control	Click through rates	Field	Immediate Post Intervention
Weaver et al. (2021)	40	Undergraduate psychology students	19	USA	e-Learning	No training control	Hits/misses/sensitivity/response bias	Lab	Immediate Post Intervention
Wen et al. (2019)	39	Volunteer university students	-	USA	Gamified (WhatHack vs. Anti-Phishing Phil vs. written material)	Between groups	% hits/misses email discrimination task; confidence; self-reported knowledge	Lab	Immediate Post Intervention
Yang et al. (2017)	63	Volunteer university students	84% 19–22, others < 30	USA	In-person training + Warnings	Warnings Only	% clicked link within phishing email; % submitting personal information	Field	3 weeks
Yeoh et al. (2022)	8,189	University staff	-	Australia	Video + e-learning vs. positive reinforcement for non-victimisation	Pre/post and Between groups	Click through rates; number of reported phishing emails	Field	
Zielinska et al. (2014)	96	Volunteers recruited online	35	USA	e-Learning + loss framing vs. e-learning + fear framing	e-Learning only	% correct email discrimination task; confidence scores	Lab	Immediate Post Intervention

Notes: i) Time of Assessment refers to the latest point at which training effects were measured post-intervention; ii) Emb = embedded condition; non-Emb = non-embedded condition; contr. = control condition; iii) Exp.1, Exp.2 = Experiment 1 / 2 respectively.

Table 4*Visual Representation of the Various Delivery Modalities and Content Formats Empirically Assessed*

Study Name	Delivery Modality						Content				
	Persistent	Standalone Passive		Standalone Interactive			Rules Based			Mindfulness Informed	Awareness
	Embedded	Text	Video / Audio	e-Learning	Game	Instructor-led	URL	Email	Social Eng.		
Abawajy (2014)		X	X		X		X	X			
Arachchilage et al. (2016)					X		X				
Back & Guerette (2021)				X							
Burns et al. (2019)	X						X	X	X		
Caputo et al. (2014)	X						X	X			X
Carella et al. (2017)	X					X	X	X	X		
Cuchta et al. (2019)	X						X	X			X
Daengsi et al. (2022)	X						X	X	X		
Davinson & Sillence (2010)					X		X				X
Dodge et al. (2012)	X						X	X			
Gokul et al. (2018)					X		X				
Gordon et al. (2019)	X						X	X	X		X
Jansson & von Solms (2013)	X										
Jensen et al. (2017)				X			X	X	X	X	X
Kavrestad et al. (2022)		X			X		X	X	X		
Kim et al. (2020)						X	X	X	X		X

Study Name	Delivery Modality						Content			
	Persistent	Standalone Passive		Standalone Interactive			Rules Based		Mindfulness Informed	Awareness
	Embedded	Text	Video / Audio	e-Learning	Game	Instructor-led	URL	Email	Social Eng.	
Kumaraguru et al. (2007a)	X						X	X		X
Kumaraguru et al. (2007b)	X	X					X	X		
Kumaraguru et al. (2008)	X						X	X	X	
Lain et al. (2021)	X						X	X		
Lim et al. (2021)		X		X			X	X		
Mayhorn & Nyeste (2012)	X				X		X			
McElwee et al. (2018)	X*									
Nguyen et al., (2021)				X			X	X	X	X
Reinheimer et al. (2020)						X	X	X		X
Roepke et al. (2022)					X		X			
Sarno et al. (2022)		X					X	X	X	X
Sharevski & Jachim (2022)			X				X	X	X	X
Sheng et al. (2010)		X			X		X	X	X	X
Silic & Lowry (2020)		X			X		X	X	X	X
Singh et al. (2023)				X			X	X	X	
Stockhardt et al. (2016)		X		X		X	X			X
Sumner et al. (2022)				X			X	X	X	X
Sutter et al (2022)	X						X	X	X	X

Study Name	Delivery Modality						Content			
	Persistent	Standalone Passive		Standalone Interactive			Rules Based		Mindfulness Informed	Awareness
	Embedded	Text	Video / Audio	e-Learning	Game	Instructor-led	URL	Email	Social Eng.	
Tschakert & Ngamsuriyaroj (2019)		X	X		X	X	X	X		X
Volkamer et al. (2018)			X				X	X	X	X
Wash & Cooper (2018)	X						X	X	X	
Weaver et al. (2021)				X			X	X		X
Wen et al. (2019)		X			X		X	X	X	
Yang et al. (2017)						X	X	X		X
Yeoh et al. (2022)	X									
Zielinska et al. (2014)				X			X	X	X	X

Notes: i) shading has been used for ease of reading purposes; ii) X*; McElwee et al. (2018) is a correlational study examining retrospective data including click-rates and phishing email simulations. No information was provided on whether training/feedback was incorporated into the simulation exercises; iii) Rules-Based training included how to identify phishing emails via URLs, email content cues (e.g., spelling/grammatical errors, sender address) or social engineering cues (e.g., calls to action, urgency); iv) Awareness training included descriptions of phishing attacks, their prevalence and potential risks; v) Mindfulness informed programs focused on raising suspicion of contextual cues (e.g., motivation of the request).

Embedded Training

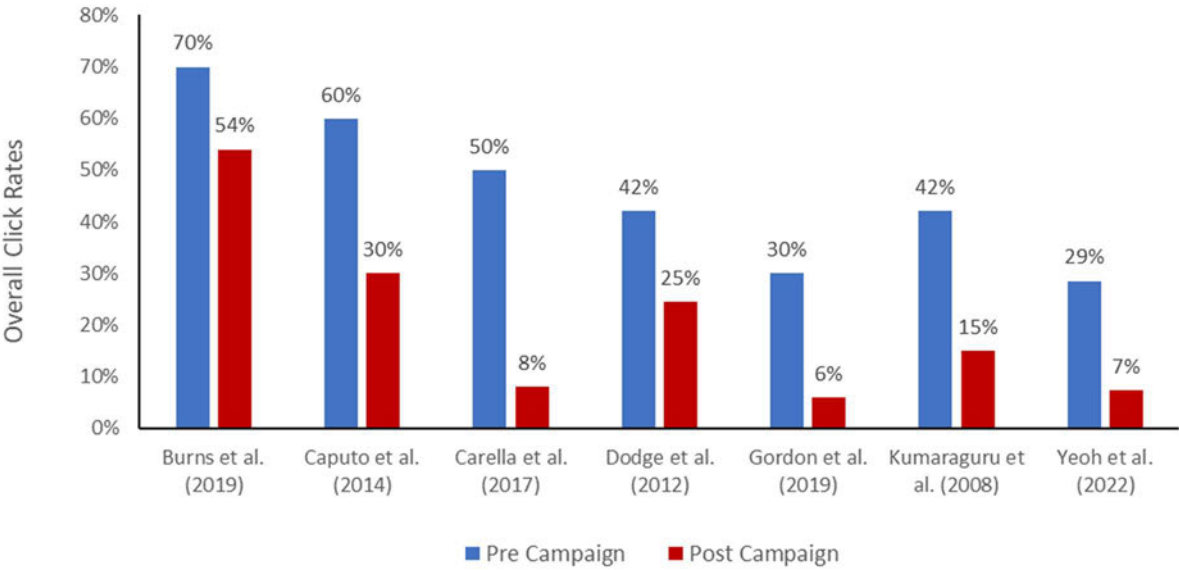
Embedded training integrates ‘teachable moments’ during regular use of email. Typically, this approach involves simulated phishing attacks wherein periodic phishing emails are sent to users. If the user responds to this phishing email (e.g., clicking on a malicious link), they receive immediate feedback including actionable steps to avoid similar attacks (Kumaraguru et al., 2007). To this end, embedded training has been described as a form of behavioural conditioning wherein reinforcement is provided at the point of susceptibility, differentiating it from other forms of standalone modalities (Caldwell, 2016).

Sixteen studies (38%) evaluated embedded training, making it the most frequent modality assessed. Of those, the majority were large-scale field-based studies ($k = 13$; 81%) involving multiple simulation rounds conducted over a period of weeks (Burns et al., 2019; Dodge et al., 2012; Jansson & von Solmns, 2013) to months (Gordon et al., 2019; Lain et al., 2021; Sutter et al., 2022; Yeoh et al., 2022). Studies typically demonstrated downward trends in click-rates over repeated simulations, with spikes exhibited with more challenging campaigns (Gordon et al., 2019; Jansson & von Solmns, 2013; Yeoh et al., 2022).

A wide distribution of pre/post click-through rates was evident (see Figure 2). Variance between study design prevents our ability to draw conclusive evidence, however, it would appear that lower rates were associated with more intensive campaigns offering a greater volume of simulations (Carella et al., 2017; Gordon et al., 2019; Yeoh et al., 2022; see Figure 3) relative to less intensive programs (Dodge et al., 2012) or those utilising targeted spear-phishing simulated attacks (Burns et al., 2019; Caputo et al., 2014).

Figure 2.

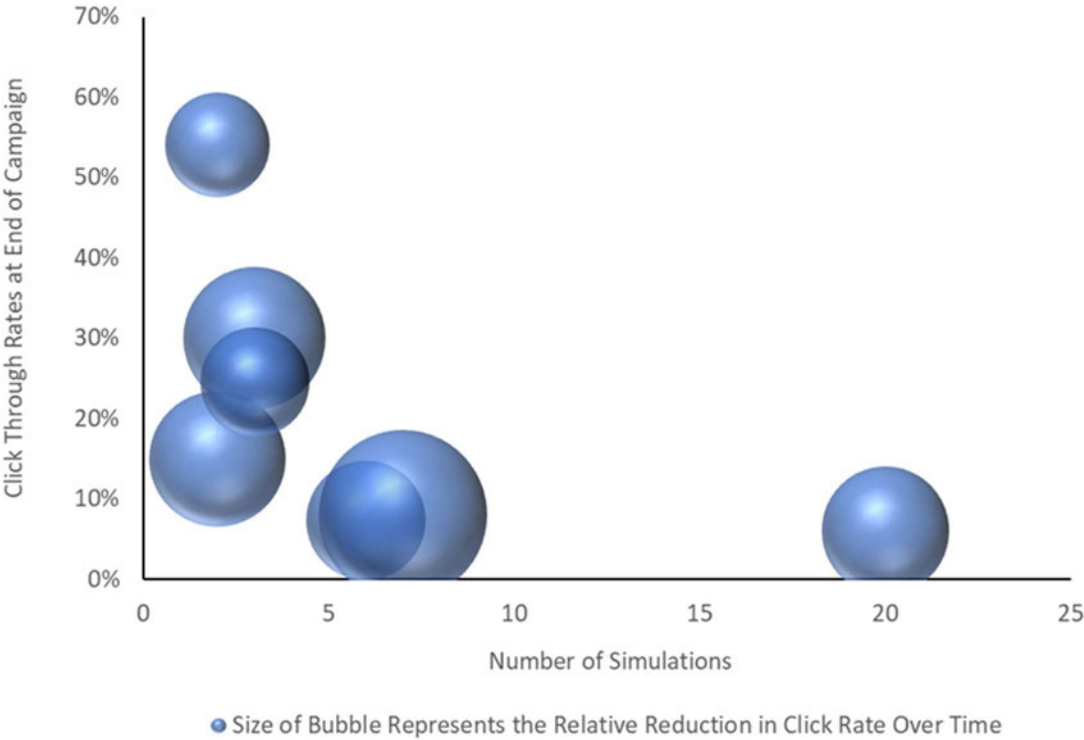
Overall Click Rates Recorded Pre and Post Embedded Training Simulation Exercises



Note. Data is provided for those studies reporting this information; Gordon et al. (2019) reflects pre/post results from those participants categorised as ‘non-offenders’

Figure 3.

Comparison of Click-Through Rates and Simulation Rounds



Five studies examined differences in click rates between repeat offenders and non-offenders and found that training significantly reduced the likelihood of future clicks (Burns et al., 2019; Caputo et al., 2014; Dodge et al., 2012; Gordon et al., 2019; Sutter et al., 2022). Although, repeat offenders remained more likely to click post-training compared to their non-trained, non-click counterparts. That is; training was able to reduce the gap between groups, but not completely.

In contrast Lain et al. (2021) found that susceptibility to phishing emails *increased* for those who received embedded training. The authors caution the use of this method which led some participants to become overly confident in the organisation's IT systems. However, succinct warnings atop suspicious emails (i.e., "looks suspicious") were effective as was the integration of a simple reporting function embedded within the email interface alongside positive feedback for its use.

Studies comparing embedded interventions to no-training-controls ($k = 11$; 69%), found a positive effect associated with training. However, the relationship was not always statistically significant ($p > .05$; Caputo et al., 2014) or in some cases was only evident after a period of delay (Gordon et al., 2019; Sutter et al., 2022). Three studies recorded a positive effect for embedded training over and above 'feedback only' conditions (i.e., notification of victimisation *without* subsequent training on identifiable cues; Dodge et al., 2012; Jansson & von Solms, 2013; Sutter et al., 2022). That is, lower click-rates were associated with programs who trained users how to identify phishing emails in addition to exposing them to simulated attacks. Despite this finding, Jansson and von Solms (2013) and Sutter et al. (2022) found that the vast majority of 'trained' participants (80% and 90% respectively) did not complete the assigned voluntary e-learning.

Several studies have evaluated embedded training against non-embedded modalities. Embedded training delivered via online documents was associated with lower click-rates when compared to in-person training covering the same material (Carella et al., 2017). Similarly, Kumaraguru et al. (2007b) compared a comic strip-styled training page either delivered to an email inbox or embedded within the simulation exercise such that it appeared immediately after clicking.

The embedded intervention significantly ($p < .01$) improved participants' ability to identify phishing emails compared to those in the non-embedded condition who performed similarly to no-training controls.

Varying formats of embedded training have also been compared to ascertain whether some modes are more effectual than others. Game formats, long-form documents, and succinct visual documents were all found to reduce click-rates to a similar degree (Cuchta et al., 2019), as were interactive, non-interactive and combined methods of embedded training (Sutter et al., 2022). In contrast, comic-strip formats (i.e., less text, more graphics, and a storyline) significantly improved email discrimination performance when compared to lengthier, fact-based modalities (Kumaraguru et al., 2007a), although Wash and Cooper (2018) found that lowest post-training click rates were associated with 'facts-and-advice' training delivered by an expert.

Standalone Training via Passive Modalities

Passive delivery describes static forms of training wherein users have the flexibility to choose the pace of learning without the benefit of feedback or other interactive elements. Our sample revealed three modalities within this category, comprising text ($k = 10$), video ($k = 1$), and audio-based ($k = 1$) approaches. Studies typically lacked detailed information on the material covered other than indicating overarching themes. These included an introduction to phishing, possible consequences, and tips on how to avoid falling victim to phishing attempts.

While 10 articles involved text-based interventions, most (60%) were as comparator conditions to more interactive primary interventions (i.e., gamification and e-learning; see Table 4). Primary interventions were associated with greater phishing resilience compared to the those delivered as text, which in most cases were described as ineffectual (Kumaraguru et al., 2007b; Lim et al., 2021; Silic & Lowry, 2020; Wen et al., 2019). Across studies there was considerable variation in the proportion of incorrect legitimacy decisions immediately post-training, from 12% (Stockhardt et al., 2016) to 40% (Silic & Lowry, 2020).

Sarno et al. (2022) developed a novel phishing classification aid consisting of seven questions to evaluate email legitimacy (e.g., ‘does the content seem unreasonable?’). The aid significantly improved phishing email discrimination compared to no training controls and feedback-only conditions, including under conditions of high email load and low phishing email prevalence. The intervention acted to lower response times which may have invoked more systematic processing, thus greater precision. However, improved classification accuracy did not consistently translate to safe actions, with participants selecting inappropriate actions (e.g., open attachment) on 20-30% of phishing emails.

Standalone Training via Interactive Modalities

Gamification. Virtual game-based modalities combine training content with goal-oriented play (Alhashmi et al., 2021). Gamification methods purport to offer an effective approach for improving user engagement, intrinsic motivation and subsequent security compliance (Sheng et al., 2007; Silic & Lowry, 2020; Wen et al., 2019). Our data revealed that games involved teaching users to distinguish phishing URLs and websites from legitimate sources (Arachchilage et al., 2016; Gokul et al., 2018; Roepke et al., 2022; Sheng et al., 2007) and how to identify malicious attachments (Silic & Lowry, 2020; Wen et al., 2019).

The vast majority of studies were lab-based experiments ($k = 9$; 82%) involving relatively small samples of volunteer undergraduate students ($k = 6$; 55%) with outcome variables measured immediately post intervention ($k = 10$; 91%). More than half ($k = 6$; 55%) of our sample involved researchers evaluating their own proprietary software (Arachchilage et al., 2016; Gokul et al., 2018; Roepke et al., 2022; Sheng et al., 2010; Silic & Lowry, 2020; Wen et al., 2019).

Several studies found that trained participants were less likely to be phished immediately post-game compared to controls (Kavrestad et al., 2022; Sheng et al., 2010; Silic & Lowry 2020). Similarly, studies utilising pre-post-test designs reported immediate improvements in users’ phishing URL detection accuracy (Arachchilage et al., 2016; Gokul et al., 2018; Roepke et al., 2022; see Table 5). The one study evaluating training retention found that users’ URL classification

performance deteriorated considerably after three months, although remained significantly better than pre-test levels (Roepke et al., 2022).

While most papers report overall group effects associated with gamification, differential training outcomes have been observed in sub-group analyses. Gokul et al. (2018) found that significant post-training improvements in URL classifications were only apparent for those with little to moderate domain-specific knowledge prior to the intervention. Silic and Lowry (2020) demonstrated that heightened trainee engagement, established through optimal challenge, incentives, and unambiguous feedback, positively influenced key outcomes.

Similar results were reported for *Anti-Phishing Phil*; a game designed in accordance with several learning science principles (i.e., procedural knowledge, reflection, and story-based content; Sheng et al., 2007). Through a series of instructional tips (e.g., ‘URLs with numbers in front are generally a scam’), the game teaches users to recognise phishing URLs and how to use search engines to find legitimate sites. While training was associated with a substantive 40% reduction in susceptibility, trainees nonetheless fell for 28% of phishing emails. In contrast, 11% of participants misclassified phishing or legitimate emails after playing ‘*What.Hack*’; a game which uses less prescriptive rules (e.g., ‘allow emails from trusted domains’ as opposed to ‘never click on links’) and includes protective strategies for content-based email phishing attacks (i.e., spear-phishing; Wen et al., 2019). Compared to *Anti-Phishing Phil* and typical ‘fact and advice’ online training materials, *What.Hack* was associated with a significant improvement (36.7%; $p < .01$) in identifying phishing emails:

Table 5*Comparison of Phishing Susceptibility Rates Pre/Post Gamification Interventions*

Study Name	Pre	Post	Measure
Arachchilage et al. (2016)	56%	84%	Mean correctness; URL discrimination
Gokul et al. (2018)	78%	86%	Correctly classified phishing URLs
Roepke et al. (2022)	70-73%	82%-84%	Mean correctness; URL discrimination
Sheng et al. (2011)	53%	72%	Proportion of participants not phished
Silic & Lowry (2020)		73%	Proportion of participants not phished
Wen et al. (2019)	65%	89%	Mean correctness; URL discrimination

Note. Data is provided for those studies reporting this information.

e-Learning. E-learning is a form of self-directed online training typically combining educational videos with knowledge tests to assess understanding, and in turn, provide trainees with direct feedback (Alhasmi et al., 2021). E-learning is recognised as a cost-effective means by which organisations are able to train employees to a corporate-wide standard (Alhasmi et al., 2021). However, for some user groups, this may come at the expense of monotonous and irrelevant content that users click through with minimal engagement (Reeves et al., 2021a; Williams et al., 2018).

Our sample revealed five studies examining e-learning as a standalone intervention (Back & Guerette, 2021; Lim et al., 2021; Stockhardt et al., 2016; Sumner et al., 2022; Weaver et al., 2021). Training typically involved completing several brief modules and passing a short test before progression to the next level was possible (Sumner et al., 2022; Stockhardt et al., 2016). Topics included a description of phishing, potential consequences of victimisation, cues for identifying phishing emails (e.g., hover over URLs), and advice on good security behaviours. Most studies ($k = 4$; 80%) involved lab-based experiments with relatively small samples of student volunteers, with training outcomes measured immediately post intervention. While training was associated with

improvements in the detection of phishing emails, on average, trained participants remained vulnerable to one in three phishing emails (see Table 6).

Table 6

Comparison of Phishing Susceptibility Rates Pre/Post e-Learning Interventions

Study Name	Pre	Post	Measure
Lim et al. (2021)	41%	48%	Mean correctness; phishing emails
Stockhardt et al. (2016)	78%	86%	Mean correctness; phishing URLs
Sumner et al. (2022)	59%	68%	Mean correctness; URL discrimination
Weaver et al. (2021)	40%	67%	Mean correctness; email discrimination

Note. Data is provided for those studies reporting this information.

An exception was found in a large-scale retrospective study wherein training had the opposite effect to that predicted (Back & Guerette; 2021). Employees who had completed a mandatory cybersecurity awareness program were six times more likely to open phishing emails, four times more likely to click on embedded links, and twice as likely to submit personal information compared to their non-trained counterparts. The quasi-experimental nature of this study, however, allows for the possibility that inherent differences between treatment and comparison groups are responsible for the contradictory findings.

For those papers reporting positive training effects for e-learning, several observations were noteworthy. Firstly, participants receiving more extensive training with opportunities for practice were significantly more resilient to phishing emails, whereas less extensive programs not only failed to facilitate phishing detection but also harmed user confidence (Lim et al., 2021). Secondly, participants who had previously encountered phishing training performed better than their previously untrained peers (Sumner et al., 2022), suggesting repeated exposure provides greater resilience. Lastly, instructor-led training was more effective than e-learning and delivering the same information in writing. Further, this method was also rated most favourably for user satisfaction despite taking longer to complete (Stockhardt et al., 2016). Notably, the instructor-led training

involved opportunities for practice, questions, elaboration, and feedback, while comparator conditions did not.

Instructor-led Training. Instructor-led modalities exist in various formats, including lecture-styled seminars and classroom-based workshops. Common to each is that training occurs in person, allowing for direct interaction and clarification. It also provides an opportunity for the instructor to customise the content and cadence of delivery to suit individual learners' needs. While users tend to express a preference for this modality (Stockhardt et al., 2016; Tschakert & Ngamsuriyaroj, 2019), as a comparatively more expensive and time-intensive medium, it is important to understand the justification for its use (Kumaraguru et al., 2008; Valentine, 2006).

Five studies within our dataset measured the effect of instructor-led training on user susceptibility to phishing emails (see Table 4). Program duration varied considerably across the sample, from 10 minutes (Yang et al., 2017) to several hours (Kim et al., 2020; Reinheimer et al., 2020). While each intervention covered similar content (i.e., general awareness of phishing, its implications, and common indicators of illegitimate emails), longer programs included more extensive examples alongside opportunities to put learning into practice.

Without exception, in-person training improved near-term user resilience to phishing emails compared to pre-training levels (see Table 7). Additionally, in comparison to delivering the same content via hardcopy material or interactively online, instructor-led training achieved the highest improvements in URL accuracy scores (Stockhardt et al., 2016).

Table 7*Comparison of Phishing Susceptibility Rates Pre/Post Instructor-led Interventions*

Study Name	Pre	Post	Delay	Measure
Carella et al. (2017)	48%	64%	50%	Proportion not phished; delay at 8 weeks
Reinheimer et al. (2020)	62% / 69%	80% / 79%	73% / 70%	Mean correctness; phish/legit emails; delay at 6 months
Stockhardt et al. (2016)	63% / 69%	91% / 97%		Mean correctness; phish/legit emails

Note. Data is provided for those studies reporting this information.

While studies measuring the endurance of training effects over time are lacking, Reinheimer et al. (2020) demonstrated that email discrimination task performance was no longer significantly different to pre-training levels six months after ‘three-to-four’ hours of instructor-led training. However, for those in the brief refresher training condition, effects were able to be sustained until at least twelve months (Reinheimer et al., 2020). In contrast, Carella et al. (2017) found that click-through rates returned to their pre-training levels within eight weeks of a brief, 20-to-30-minute, instructor-led intervention.

Anti-Phishing Training Content and Pedagogical Approach

Studies generally provided limited information on training content, and in some cases, this information was absent altogether (Back & Guerette, 2021; Jansson & von Solms, 2013; Mayhorn & Nyeste, 2012; McElwee et al., 2018; Yeoh et al., 2022). Where reported, most interventions involved training users to identify common indicators of phishing emails (i.e., rules-based training). Alternate pedagogical approaches included various message framing conditions as motivators of behavioural change (e.g., Davinson & Sillence, 2010), various feedback conditions and frequencies of phishing emails (e.g., Singh et al., 2023), and ‘mindful awareness training’ as an alternative to traditional rules-based methods (Jensen et al., 2017; Nguyen et al., 2021).

Rules-Based Training The primary emphasis of most interventions was on explaining common phishing methods and educating users how to identify phishing emails. This was achieved through rule-based instructions (e.g., hover over links to explore URLs; Caputo et al., 2014; Kumaraguru et al., 2008; Sheng et al., 2010) and plausibility checks (e.g., grammatical/spelling errors; Lim et al., 2021; Reinheimer et al., 2020; Sarno et al., 2022; Sutter et al., 2022). The rules-based approach to training teaches users to habitually apply learned behaviours, specifically, to identify common cues of phishing emails and then apply protective strategies. Widely studied, empirical evidence has confirmed that this method increases user resistance to phishing attacks (Kumaraguru et al., 2010; Sheng et al., 2010).

Framing Effects as Persuasive Appeals to Enhance Training Outcomes

Several studies measured the impact of heightened fear or loss messaging within their campaigns by emphasising the potential negative consequences of phishing victimisation (Burns et al., 2019; Caputo et al., 2014; Zielinska et al., 2014). Despite fear and loss conditions performing marginally better than comparators, group differences were not significant. Specifically, those in message framing conditions performed similarly to typical ‘fact and advice’ trainees (Zielinska et al., 2014) and no training controls (Caputo et al., 2014); including whether the loss was framed as an individual or group-level impact (Burns et al., 2019). Similarly, Davinson and Sillence (2010) manipulated user susceptibility and user control to explore their effects on intended and actual secure behaviours. No significant differences were observed between those in high- or low-susceptibility conditions, irrespective of training (i.e., coping control). Participants across each of the conditions equally intended to behave more securely than they had indicated at baseline, yet at the 7-day follow-up stage self-reported significantly less secure behaviours than anticipated.

Experiential Learning

Providing trainees with opportunities to ‘learn by doing’ (i.e., experiential learning), through repeated practice and performance feedback, was recognised as an important aspect of training design (Reinheimer et al., 2020; Singh et al., 2023; Stockhardt et al., 2016;). However, Singh et al.

(2023) found that a higher frequency of phishing emails during training was associated with a higher propensity for classifying subsequent emails as phishing. That is, trainees became more cautious without an improvement in discriminability (i.e., more hits *and* false alarms). Although, providing detailed feedback on phishing cues significantly ($p < .001$) improved classification accuracy under frequency conditions, compared to outcome-based feedback (i.e., ‘correct/incorrect’).

Mindful Awareness Training

Recently, researchers have developed a novel approach to anti-phishing training by incorporating mindfulness principles within their programs (Jensen et al., 2017; Ngyuen et al., 2021). Designed as a supplement to rules-based training, this method teaches users to mindfully attend to message evaluations, particularly to the context within which messages are received. Thought to encourage more systematic processing, this method asks users to pause any time an email requires action, to consider the context and possible motivation for the request, and where suspicion is aroused, to check with a trusted third party (Jensen et al., 2017).

In comparison to rules-based training, mindfulness-trained participants were significantly less likely to respond to phishing emails (Jensen et al., 2017). Delivered as a brief online training format, the authors demonstrated that the mindfulness approach could be used as relatively simple yet effective supplement to rules-based training, particularly for individuals with less experience and education, and those low in email mindfulness. Similar results were confirmed by Ngyuen et al. (2021) who demonstrated that compared to no-training and rules-based training comparators, the mindfulness approach significantly improved email discrimination accuracy, without a subsequent increase in false positives. Moreover, Ngyuen et al. (2021) showed that these effects persist for at least two months post training.

Discussion

The aim of this scoping review was to identify and describe the nature of available anti-phishing training methods and their measurable outcomes on user susceptibility to phishing emails.

We present our discussion by summarising general findings on the status of the evidence base before addressing key findings specific to each modality. We then discuss in detail emergent themes that speak to common elements of effective training design across studies. Finally, study limitations and directions for future research are addressed.

General Summary of Findings

User-training is promoted as an important and effective mitigant against the human factors associated with phishing email susceptibility (Abawajy, 2012; Dodge et al., 2012; Kumaraguru et al., 2010). However, our study has revealed that the evidence supporting these claims is relatively immature. Although considerable contributions have been made by researchers in the field, the majority of studies evaluating anti-phishing training methods have tended to focus on short-term outcomes, with effectiveness evaluated against no-training controls, or through pre/post-test comparisons. Given the inherent limitations associated with these methodological approaches (Mann, 2003), the evidence for interventions achieving their intended objectives of training users to identify phishing messages, and apply protective actions consistently over time, is limited.

Additionally, many studies ($k = 52\%$) have relied on tests of knowledge to measure training efficacy (e.g., email legitimacy tasks, URL classifications) under contrived experimental conditions that artificially raise phishing vigilance as the primary goal (Parsons et al., 2013, Parsons et al., 2015). While we can observe that knowledge has been acquired, these measures do not provide real insight on the success of programs in driving sustained behavioural change under naturalistic conditions (e.g., time pressure and high email load; Butavicius et al., 2022; Jones et al., 2015). To do so requires measures of training impact that take into account knowledge acquisition, recollection, and transfer under realistic experimental manipulations (Baldwin & Ford, 1988; Shoeb, 2019). While not perfect, click-rates observed during simulation studies represent a more valid measure of susceptibility and thus a reasonably good indication of learning transfer (Jones et al., 2015). Such methods were more consistently applied in studies evaluating embedded training methods for which the evidence base is more advanced.

A number of different approaches for delivering training have been evaluated to varying degrees. These include persistent methods (i.e., embedded training) or standalone campaigns involving active participation (i.e., e-learning, gamification, instructor-led programs) or passive knowledge acquisition (i.e., text, videos). Across modalities, training improved user resilience to phishing emails, confirming its place as an important defensive mechanism. Statistically significant differences were observed between trained and untrained participants, along with improvements in rates of phishing email detection compared to baseline measures (e.g., Sheng et al., 2010; Sumner et al., 2022; Weaver et al., 2021). However, consistent with prior research (Alhashmi et al., 2021), on average, 23% of users continued to be deceived post training. While cybersecurity experts caution that the evolving nature of phishing attacks and their increasing sophistication likely prohibits a zero click-through rate (Greene et al., 2018), the question remains whether 23% is an acceptable level of residual risk. At a minimum it suggests that current approaches to training have room to improve.

Evidence for near-term training effects are well documented. Many studies reported positive results within 7 days of training when acquired knowledge and skills are easily retrieved (e.g., Stockhardt et al., 2016; Wen et al., 2019). However, measures of training retention beyond this term were limited. A mere five studies (12%) considered follow-up periods of eight weeks or more (Back & Guerette, 2021; Gordon et al., 2019; Kim et al., 2020; Reinheimer et al., 2020; Roepke et al., 2022). As evidenced within the training literature, effects typically diminish with learning decay as the time since training becomes more distal (Blume et al., 2010). Correspondingly, studies utilising delayed outcome measures to verify training maintenance reported significant deterioration in phishing resilience over time. Our dataset revealed that, at most, training effects were able to be sustained up to six months; although providing brief refresher material to remind users of the training extended this to 12 months (Reinheimer et al., 2020). Findings therefore suggest that annualised programs, common to many organisations, should be reconsidered (ProofPoint, 2020).

Training Modalities Empirically Evaluated

Embedded training was the most widely researched modality. This method has been shown to substantively reduce user susceptibility to phishing emails over multiple simulation rounds (Carella et al., 2017; Dodge et al., 2012; Gordon et al., 2019; Yeoh et al., 2022). Studies found that embedding training at the point of ‘dangerous clicking’ was more effectual than delivering it separately (e.g., via emails or games; Kumaraguru et al., 2007b), supporting the view that individuals are more receptive to training at the point of error, when feedback is timely and directly relevant (Salas & Cannon-Bowers, 2001).

Across organisations, e-learning is one of the most frequently used methods to deliver cybersecurity training (ProofPoint, 2023). However, our results suggest that the evidence for e-learning’s efficacy in reducing user susceptibility to phishing emails is limited, with no empirical validation of behavioural impact beyond the immediate term, and trained participants on average remaining vulnerable to one in three phishing emails (Sumner et al., 2022; Weaver et al., 2021), or in some cases more exposed (Back & Guerette, 2021). Moreover, the majority of trainees did not complete the optional training provided (Jansson & von Solmns, 2013; Sutter et al., 2022). Such outcomes lend support to the view that online methods tend to be unengaging, out of context, and allow trainees to quickly click through content without absorbing it (Caldwell, 2016; Reeves et al., 2021a; Williams et al., 2018).

Instructor-led training was commonly the modality preferred by users, consistent with prior research (Reeves et al., 2021a). When implemented intensively over several hours, it was found to significantly reduce users’ phishing susceptibility (Kim et al., 2020; Reinheimer et al., 2020). However, as a comparatively more expensive method, evidence suggests instructor-led training brings little substantive value over other self-paced methods (Tschakert & Ngamsuriyaraj, 2019), although this study was significantly underpowered to detect small to moderate effects (Cohen, 1992).

While an increasing number of game-based cybersecurity products have been commercialised, rigorous evidence validating their use has previously been described as lacking (Tioh et al., 2017). Our findings tend to support this view. Gamified approaches have focused on training URL discrimination for which improvements have been demonstrated in the near term, particularly for those with little prior domain-specific knowledge (Gokul et al., 2018). Results suggest that gamification may be a novel method for upskilling novices for a very specific purpose (i.e., URL knowledge; Arachchilage et al., 2016; Gokul et al., 2018; Roepke et al., 2022; Sheng et al., 2010; Wen et al., 2019), although the evidence-base lacks sufficient longitudinal studies that assess learning retention and application over time.

Features of Training Delivery Associated with Improved Outcomes

While each modality involved similar instructional content, certain delivery methods provided additional features that facilitated enhanced engagement that not all techniques could support. This review found evidence that highly engaging methods that provided opportunities to practice learned skills and receive feedback were more effective than passive forms of training (e.g., emailed materials). This corresponds with extensive evidence in the training literature that active approaches to learning are superior for knowledge retention and the transfer of learned skills ‘beyond the classroom’ (Burke & Hutchins, 2007; Grossman & Salas, 2011; Salas & Cannon-Bowers, 2001). However, when low-engagement methods were used as a reinforcement strategy for more active approaches, positive responses were observed (Abawajy, 2014; Lim et al., 2021; Reinheimer et al., 2020). As skilled behaviour is thought to develop through the accumulation of repeated practice and exposure (Rasmussen, 1986), it follows that programs providing increased opportunities *for* practice and exposure to a variety of phishing emails may be more efficacious. Specifically, for the development of a large and varied repertoire of experiences, or ‘mental models’ from which to draw from (Phillips et al., 2004).

In addition, feedback proved especially important for the calibration of confidence and performance in legitimacy decisions (Lim et al., 2021; Singh et al., 2023). Similar trends are

observed in other domains of expertise involving visual search tasks with a low base rate of signals such as baggage screening and medical diagnosis (Evans et al., 2013; Wolf et al., 2013).

Importantly, Singh et al. (2023) found that increased frequency of phishing emails in the absence of detailed feedback led to higher false positives and over-confidence. However, the combination of exposure *and* detailed feedback (i.e., describing the cues leading to the outcome, rather than outcome-based, ‘correct/incorrect’ feedback) significantly improved email discrimination accuracy. Findings lend support to the importance of compiling an extensive experience bank alongside accurate, timely, and diagnostic feedback (Phillips et al., 2004). Specifically, the provision of process-based feedback which seeks to strengthen learning by informing people of the necessary changes to their approach, as opposed to merely indicating whether they are improving or not (Salas & Cannon-Bowers, 2001). From a practical standpoint, specificity of feedback and high base rates of exposure to phishing emails should be strongly considered within training design.

The frequency and length of training interventions varied considerably across studies. Controlling for modality type, more intensive programs were generally associated with more positive training outcomes; a trend similarly observed in the security awareness literature (Kweon et al., 2021). For example, a brief tutorial reduced click-through rates to 36% although effects were short-lived (Caraella et al., 2017), whereas a three-to-four-hour tutorial covering similar content reduced click-rates to 20% and was able to be sustained for at least six months (Reinheimer et al., 2020). While results indicate support for lengthier programs, we note that the feasibility of the training protocol by Reinheimer et al. (2020) would be challenging to implement practically. There is also the need to establish the ‘right’ volume of training to maximise attentional awareness without overwhelming users such that they disengage (Reeves et al., 2021b). This is particularly relevant in the context of significant governance and compliance obligations that equally compete for employees’ attention (Alshaikh, 2018). Future empirical explorations would benefit from the consideration of the interactions between these variables, to potentially identify an inflection point wherein more training becomes harmful to outcomes. At the same time, we acknowledge that

training intensity is highly correlated with the approaches that facilitate the development of a broader and more refined set of heuristic processes (i.e., through repeated practice, exposure, and feedback), representing a challenging dynamic.

The Influence of Training Content and Pedagogical Approach

Rules-based training was the most common method of content delivery evaluated across studies. This involved teaching users common indicators of phishing emails, (e.g., spelling/grammatical errors, urgency cues) and the application of several protective strategies, or ‘rules of thumb’ to check the legitimacy of emails (e.g., manually checking URLs). Studies utilising these methods reported improved user resilience to phishing emails (e.g., Kumaraguru et al., 2007a; Sheng et al., 2010), consistent with prior research demonstrating that the skilled use of visual cues is associated with improved user discrimination of phishing emails over and above phishing-related knowledge (Bayl-Smith et al., 2020; Sturman et al., 2023). However, leveraging skilled intuition requires an accumulation of domain-specific experiences from which to draw from, acquired through repeated exposure over time (Klein et al., 1986). From the information available, we cannot infer that such skills were indeed embedded and then activated, or if improvements reflect the mere acquisition of knowledge-based ‘if-then’ rules (Rassmusen, 1986). Although, this may at least partially explain observed differences in training retention and susceptibility rates between less intensive programs (e.g., Sheng et al., 2010; Weaver et al., 2021) and those involving greater exposure to a variety of phishing emails (e.g., Carella et al., 2017; Reinheimer et al., 2020, Roepke et al., 2022; Sutter et al., 2022). As cue utilisation is thought to enable the rapid assessment of key features, it is reasoned that this method is particularly advantageous in time pressured situations (Bayl-Smith et al., 2020). Future research should therefore contrast the efficacy of targeted cues-based training against traditional knowledge-based approaches to examine this issue further.

While cue-based knowledge is a necessary pre-requisite for the ability to identify phishing emails and then respond accordingly, this method does not always translate to secure behaviours for a variety of reasons (Butavicius et al., 2022; Downs et al., 2006; Ndibwile et al., 2019). Being static

in nature, heuristic cues risk becoming quickly outdated in changeable environments, and indeed many of the earlier papers within our sample relied on cues that are no longer representative of the kinds of sophisticated phishing emails users now encounter (Alkhalil et al 2021; Ghazi-Tehrani & Pontell, 2021). Contemporary phishing attempts are less likely to contain superficial spelling/grammatical cues, and individuals must instead rely on their ability to detect more subtle signals such as the likelihood a company would ask for the details requested (Jones et al., 2015). To this end, ‘mindfulness-informed’ approaches, which ask users to stop and consider the context within which the email has been received, have recently been developed to supplement traditional rules-based training (Jensen et al., 2017). This method is thought to encourage more systematic, or deliberate, processing of email legitimacy judgements and thus greater accuracy (Luo et al., 2013). Findings revealed that in comparison to rules-based training, mindfulness-trained participants were indeed significantly less likely to respond to phishing emails (Jensen et al., 2017; Ngyuen et al., 2021). Results corroborate prior research demonstrating that individuals who are more attentive to message cues and invest more cognitive effort are less susceptible to phishing emails (Ackerley et al., 2022; Luo et al., 2013; Pattinson et al., 2012; Vishwanath et al., 2011). In practical terms this suggests that once individuals become proficient in the identification of salient features of phishing emails (i.e., cue-based knowledge), interventions would be better focused on strategies that support systematic thinking. Such findings are encouraging and warrant further investigation to see if similar effects are observed in broader populations.

Limitations and Future Directions

The diverse nature of the training interventions and study designs made it challenging to identify the specific components of training design associated with success. Additionally, some studies failed to give a detailed description of the training content while others merely listed the topics covered, thus limiting the scope of our intended review. Despite these limitations, we have nonetheless identified some common components across interventions that led to positive

outcomes. At the same time, we acknowledge that testing variable interventions provides valuable information for future studies.

The current premature stage of development of this field of research (i.e., lack of consistent, adequately powered, high-quality evidence) justifies our decision to proceed with a scoping review and has highlighted the need for more effective and standardised evaluations of training interventions intended to reduce user susceptibility to phishing emails. A potential first stage could involve establishing validated outcome measures that assess not only learning but also alterations in behaviour and the interaction between these two outcomes. The evidence base also calls for studies with greater ecological validity to understand user behaviour in real world contexts when interacting with potentially fraudulent emails, alongside the consideration of longitudinal effects to understand at what point training effects wane.

Additionally, our research did not consider the features of the mock phishing emails used across interventions and their impact on training outcomes. Emails varied in context (targeted vs. generic), alleged sender (internal vs. external), and use of persuasive techniques. Prior research has demonstrated that users are more susceptible to phishing emails from known sources (Halevi et al., 2015; Jagatic et al., 2007), whether the email is contextually relevant (Jaeger et al., 2021) and the influence techniques used (Butavicius et al., 2015; Jaeger et al., 2021; Lin et al., 2019; Vishwanath et al., 2011). It would be useful for future research to examine how various training strategies compare across various email conditions. Specifically, whether some interventions outperform others for training generalisation.

Conclusion

This scoping review explored the evidence for the efficacy of training interventions aimed at reducing end-user susceptibility to phishing emails. Components of training design that led to positive outcomes included training intensity, active approaches to learning, the provision of detailed feedback, and supplementing attentional awareness skills-based training with traditional cue-based approaches. Participants' improved knowledge and confidence in identifying phishing

emails is encouraging, although current approaches to training leave approximately 20% of users at risk. This review highlights the need for more robust studies in this area and particularly those that seek to address the cognitive biases that phishers commonly exploit. Findings provide useful insight towards the development of optimal training strategies to defend against phishing email susceptibility.

References

- Abawajy, J. (2014). User preference of cyber security awareness delivery methods. *Behaviour & Information Technology*, 33(3), 237-248. <https://doi.org/10.1080/0144929X.2012.708787>
- Ackerley, M., Morrison, B. W., Ingre, K., Wiggins, M. W., Bayl-Smith, P., & Morrison, N. (2022). Errors, irregularities, and misdirection: Cue utilisation and cognitive reflection in the diagnosis of phishing emails. *Australasian Journal of Information Systems*, 26. 1-21. <https://doi.org/10.3127/ajis.v26i0.3615>
- Albrechtsen, E., & Hovden, J. (2010). Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study. *Computers & Security*, 29(4), 432-445. <https://doi.org/10.1016/j.cose.2009.12.005>
- Aleroud, A., & Zhou, L. (2017). Phishing environments, techniques, and countermeasures: A survey. *Computers & Security*, 68, 160-196. <https://doi.org/10.1016/j.cose.2017.04.006>
- Alhashmi, A. A., Darem, A., & Abawajy, J. (2021). Taxonomy of cybersecurity awareness delivery methods: A countermeasure for phishing threats. 12(10), 29-35. *International Journal of Advanced Computer Science and Applications*. <https://doi.org/10.14569/IJACSA.2021.0121004>
- Alkhalil, Z., Hewage, C., Nawaf, L., & Khan, I. (2021). Phishing attacks: A recent comprehensive study and a new anatomy. *Frontiers in Computer Science*, 3, 563060. 1-23. <https://doi.org/10.3389/fcomp.2021.563060>
- Alshaikh, M., Maynard, S. B., Ahmad, A., & Chang, S. (2018). An exploratory study of current information security training and awareness practices in organizations. In *Proceedings of the 51st Hawaii International Conference on System Sciences*. 5085-5094. <https://doi.org/10.24251/hicss.2018.635>
- Anandpara, V., Dingman, A., Jakobsson, M., Liu, D., & Roinestad, H. (2007). Phishing IQ tests measure fear, not ability. In *Declarative Agent Languages and Technologies X*. 362–366. https://doi.org/10.1007/978-3-540-77366-5_33

- Anderson, R., Barton, C., Böhme, R., Clayton, R., Van Eeten, M. J., Levi, M., ... & Savage, S. (2013). Measuring the cost of cybercrime. *The Economics of Information Security and Privacy*, 265-300. <https://doi.org/10.17863/CAM.41598>
- Anti-Phishing Working Group. (2022). *Phishing activity trends report 4Q/2022: Unifying the global response to cybercrime*. <https://apwg.org/trendsreports/>
- Arachchilage, N. A. G., Love, S., & Beznosov, K. (2016). Phishing threat avoidance behaviour: An empirical investigation. *Computers in Human Behavior*, 60, 185-197. <https://doi.org/10.1016/j.chb.2016.02.065>
- Back, S., & Guerette, R. T. (2021). Cyber place management and crime prevention: The effectiveness of cybersecurity awareness training against phishing attacks. *Journal of Contemporary Criminal Justice*, 37(3), 427-451. <https://doi.org/10.1177/10439862211001628>
- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 41(1), 63-105. <https://doi.org/10.1111/j.1744-6570.1988.tb00632.x>
- Bayl-Smith, P., Sturman, D., & Wiggins, M. (2020). Cue utilization, phishing feature and phishing email detection. In: *Bernhard, M., et al. Financial Cryptography and Data Security. FC 2020. Lecture Notes in Computer Science*, 12063. 56-70. https://doi.org/10.1007/978-3-030-54455-3_5
- Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, 36(4), 1065-1105. <https://doi.org/10.1177/0149206309352880>
- Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review*, 6(3), 263-296. <https://doi.org/10.1177/1534484307303035>
- Burns, A. J., Johnson, M. E., & Caputo, D. D. (2019). Spear phishing in a barrel: Insights from a targeted phishing campaign. *Journal of Organizational Computing and Electronic Commerce*, 29(1), 24-39. <https://doi.org/10.1080/10919392.2019.1552745>

- Butavicius, M., Parsons, K., Pattinson, M., & McCormac, A. (2016). Breaching the human firewall: Social engineering in phishing and spear-phishing emails. *ACIS 2015 Proceedings*, 98, 1-10.
<https://aisel.aisnet.org/acis2015/98>
- Butavicius, M., Taib, R., & Han, S. J. (2022). Why people keep falling for phishing scams: The effects of time pressure and deception cues on the detection of phishing emails. *Computers & Security*, 123, 1-10. 102937. <https://doi.org/10.1016/j.cose.2022.102937>
- Caldwell, T. (2016). Making security awareness training work. *Computer Fraud & Security*, 2016(6), 8-14. [https://doi.org/10.1016/S1361-3723\(15\)30046-4](https://doi.org/10.1016/S1361-3723(15)30046-4)
- Caputo, D. D., Pfleeger, S. L., Freeman, J. D., & Johnson, M. E. (2014). Going spear phishing: Exploring embedded training and awareness. *IEEE Security & Privacy*, 12(1), 28-38.
<https://doi.org/10.1109/MSP.2013.106>
- Carella, A., Kotsoev, M., & Truta, T.M. (2017). Impact of security awareness training on phishing click-through rates. *2017 IEEE International Conference on Big Data (Big Data)*, 4458-4466.
<https://doi.org/10.1109/BigData.2017.8258485>
- Cloudian. 2021. *Cloudian ransomware survey finds 65% of victims penetrated by phishing had conducted anti-phishing training*. <https://cloudian.com/press/cloudian-ransomware-survey-finds-65-percent-of-victims-penetrated-by-phishing-had-conducted-anti-phishing-training>
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1(3), 98-101. <https://doi.org/10.1111/1467-8721.ep10768783>
- Cram, W. A., D'arcy, J., & Proudfoot, J. G. (2019). Seeing the forest and the trees: a meta-analysis of the antecedents to information security policy compliance. *MIS quarterly*, 43(2), 525-554.
<https://doi.org/10.25300/MISQ/2019/15117>
- Cuchta, T., Blackwood, B., Devine, T. R., Niichel, R. J., Daniels, K. M., Lutjens, C. H., ... & Stephenson, R. J. (2019). Human risk factors in cybersecurity. In *Proceedings of the 20th Annual SIG Conference on Information Technology Education*. 87-92.
<https://doi.org/10.1145/3349266.3351407>

- Daengsi, T., Pornpongtechavanich, P. & Wuttidittachotti, P. (2022). Cybersecurity awareness enhancement: A study of the effects of age and gender of Thai employees associated with phishing attacks. *Education and Information Technologies* 27, 4729–4752.
<https://doi.org/10.1007/s10639-021-10806-7>
- Davinson, N., & Sillence, E. (2010). It won't happen to me: Promoting secure behaviour among internet users. *Computers in Human Behavior*, 26(6), 1739-1747.
<https://doi.org/10.1016/j.chb.2010.06.023>
- Desolda, G., Ferro, L. S., Marrella, A., Catarci, T., & Costabile, M. F. (2021). Human factors in phishing attacks: a systematic literature review. *ACM Computing Surveys (CSUR)*, 54(8), 1-35. <https://doi.org/10.1145/3469886>
- Diaz, A., Sherman, A. T., & Joshi, A. (2020). Phishing in an academic community: A study of user susceptibility and behavior. *Cryptologia*, 44(1), 53-67.
<https://doi.org/10.1080/01611194.2019.1623343>
- Dodge, R.C., Coronges, K., & Rovira, E. (2012). Empirical Benefits of Training to Phishing Susceptibility. *IFIP International Information Security Conference*. 457-464.
https://doi.org/10.1007/978-3-642-30436-1_37
- Downs, J. S., Holbrook, M. B., & Cranor, L. F. (2006). Decision strategies and susceptibility to phishing. In *Proceedings of the Second Symposium on Usable Privacy and Security*. 79-90.
<https://doi.org/10.1145/1143120.1143131>
- Evans, K. K., Birdwell, R. L., & Wolfe, J. M. (2013). If you Don't find it often, you often don't find it: Why some cancers are missed in breast cancer screening. *PLoS One*, 8(5), 1–6.
<https://doi.org/10.1371/journal.pone.0064366>
- Ferguson, A. J. (2005). Fostering e-mail security awareness: The West Point carronade. *Educause Quarterly*, 28(1), 54-57.
- Furnell, S., & Clarke, N. (2012). Power to the people? The evolving recognition of human aspects of security. *Computers & Security*, 31(8), 983-988. <https://doi.org/10.1016/j.cose.2012.08.004>

- Ghazi-Tehrani, A. K., & Pontell, H. N. (2022). Phishing evolves: Analyzing the enduring cybercrime. *Victims & Offenders*, 16(3), 316-342.
<https://doi.org/10.1080/15564886.2020.1829224>
- Gokul, C.J., Pandit, S., Vaddepalli, S., Tupsamudre, H., Banahatti, V., & Lodha, S.P. (2018). PHISHY - A Serious Game to Train Enterprise Users on Phishing Awareness. *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. <https://doi.org/10.1145/3270316.3273042>
- Gordon, W. J., Wright, A., Glynn, R. J., Kadakia, J., Mazzone, C., Leinbach, E., & Landman, A. (2019). Evaluation of a mandatory phishing training program for high-risk employees at a US healthcare system. *Journal of the American Medical Informatics Association*, 26(6), 547-552.
<https://doi.org/10.1093/jamia/ocz005>
- Greene, K., Steves, M., & Theofanos, M. (2018). No phishing beyond this point. *Computer*, 51. 1-6.
<https://doi.org/10.1109/MC.2018.2701632>
- Grossman, R., & Salas, E. (2011). The transfer of training: what really matters. *International Journal of Training and Development*, 15(2), 103-120. <https://doi.org/10.1111/j.1468-2419.2011.00373.x>
- Gupta, B. B., Tewari, A., Jain, A. K., & Agrawal, D. P. (2017). Fighting against phishing attacks: state of the art and future challenges. *Neural Computing and Applications*, 28, 3629-3654.
<https://doi.org/10.1007/s00521-016-2275-y>
- Halevi, T., Memon, N., & Nov, O. (2015). Spear-phishing in the wild: A real-world study of personality, phishing self-efficacy and vulnerability to spear-phishing attacks. *Innovation Law & Policy eJournal*. <https://doi.org/10.2139/ssrn.2544742>
- Heartfield, R., Loukas, G., & Gan, D. (2016). You are probably not the weakest link: Towards practical prediction of susceptibility to semantic social engineering attacks. *IEEE Access*, 4, 6910-6928. <https://doi.org/10.1109/ACCESS.2016.2616285>

- Hu, S., Hsu, C., & Zhou, Z. (2022). Security education, training, and awareness programs: Literature review. *Journal of Computer Information Systems*, 62(4), 752-764.
<https://doi.org/10.1080/08874417.2021.1913671>
- IBM Security (2023). *Cost of a Data Breach Report 2023*.
<https://www.ibm.com/downloads/cas/E3G5JMBP>
- Jaeger, L., & Eckhardt, A. (2021). Eyes wide open: The role of situational information security awareness for security-related behaviour. *Information Systems Journal*, 31(3), 429-472.
<https://doi.org/10.1111/isj.12317>
- 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>
- Jakobsson, M. (2007). The human factor in phishing. *Privacy & Security of Consumer Information*, 7(1), 1-19. <https://doi.org/10.4324/9780203257487>
- Jampen, D., Gür, G., Sutter, T., & Tellenbach, B. (2020). Don't click: towards an effective anti-phishing training. A comparative literature review. *Human-centric Computing and Information Sciences*, 10(1), 1-41. <https://doi.org/10.1186/s13673-020-00237-7>
- Jansson, K., & von Solms, R. (2013). Phishing for phishing awareness. *Behaviour & Information Technology*, 32(6), 584-593. <https://doi.org/10.1080/0144929X.2011.632650>
- Jensen, M. L., Dinger, M., Wright, R. T., & Thatcher, J. B. (2017). Training to mitigate phishing attacks using mindfulness techniques. *Journal of Management Information Systems*, 34(2), 597-626. <https://doi.org/10.1080/07421222.2017.1334499>
- Jones, H. S., Towse, J. N., & Race, N. (2015). Susceptibility to email fraud: A review of psychological perspectives, data-collection methods, and ethical considerations. *International Journal of Cyber Behavior, Psychology and Learning*, 5(3), 13-29.
<https://doi.org/10.4018/IJCBPL.2015070102>
- Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6), 515-526. <https://doi.org/10.1037/a0016755>

- Kävrestad, J., Hagberg, A., Nohlberg, M., Rambusch, J., Roos, R., & Furnell, S. (2022). Evaluation of contextual and game-based training for phishing detection. *Future Internet*, 14(104), 1-16.
<https://doi.org/10.3390/fi14040104>
- Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing detection: a literature survey. *IEEE Communications Surveys & Tutorials*, 15(4), 2091-2121.
<https://doi.org/10.1109/SURV.2013.032213.00009>
- Kim, B., Lee, D. Y., & Kim, B. (2020). Deterrent effects of punishment and training on insider security threats: a field experiment on phishing attacks. *Behaviour & Information Technology*, 39(11), 1156-1175.
- Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid Decision making on the Fire Ground. *Proceedings of the Human Factors Society Annual Meeting*, 30(6), 576-580.
<https://doi.org/10.1177/154193128603000616>
- Klein, G. A., & Calderwood, R. (1991). Decision models: Some lessons from the field. *IEEE Transactions on Systems, Man, and Cybernetics*, 21(5), 1018-1026.
<https://doi.org/10.1109/21.120054>
- Kumaraguru, P., Rhee, Y., Acquisti, A., Cranor, L. F., Hong, J., & Nunge, E. (2007a). Protecting people from phishing: the design and evaluation of an embedded training email system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 905-914.
<https://doi.org/10.1145/1240624.1240760>
- Kumaraguru, P., Rhee, Y., Sheng, S., Hasan, S., Acquisti, A., Cranor, L. F., & Hong, J. (2007b). Getting users to pay attention to anti-phishing education: evaluation of retention and transfer. In *Proceedings of the Anti-Phishing Working Groups 2nd Annual eCrime Researchers Summit* (70-81). <https://doi.org/10.1145/1299015.1299022>
- Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2008). Lessons from a real-world evaluation of anti-phishing training. In *2008 eCrime Researchers Summit*. 1-12.
<https://doi.org/10.1109/ECRIME.2008.4696970>

- Kumaraguru, P., Sheng, S., Acquisti, A., Cranor, L. F., & Hong, J. (2010). Teaching Johnny not to fall for phish. *ACM Transactions on Internet Technology (TOIT)*, 10(2), 1-31.
<https://doi.org/10.1145/1754393.1754396>
- Kweon, E., Lee, H., Chai, S., & Yoo, K. (2021). The utility of information security training and education on cybersecurity incidents: An empirical evidence. *Information Systems Frontiers*, 23, 361-373. <https://doi.org/10.1007/s10796-019-09977-z>
- Lain, D., Kostiainen, K., & Čapkun, S. (2022). Phishing in organizations: Findings from a large-scale and long-term study. In *2022 IEEE Symposium on Security and Privacy (SP)*. 842-859.
<https://doi.org/10.1109/SP46214.2022.9833766>
- Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation Science*, 5, 1-9. <https://doi.org/10.1186/1748-5908-5-69>
- Lim, J., Zhou, L., & Zhang, D. (2021). Verbal deception cue training for the detection of phishing emails. In *2021 IEEE International Conference on Intelligence and Security Informatics (ISI)*. 1-3. <https://doi.org/10.1109/ISI53945.2021.9624738>
- Lin, T., Capecci, D. E., Ellis, D. M., Rocha, H. A., Dommaraju, S., Oliveira, D. S., & Ebner, N. C. (2019). Susceptibility to spear-phishing emails: Effects of internet user demographics and email content. *ACM Transactions on Computer-Human Interaction: A Publication of the Association for Computing Machinery*, 26(5), 1-28. <https://doi.org/10.1145/3336141>
- Luo, X. R., Zhang, W., Burd, S., & Seazzu, A. (2013). Investigating phishing victimization with the Heuristic–Systematic Model: A theoretical framework and an exploration. *Computers & Security*, 38, 28-38. <https://doi.org/10.1016/j.cose.2012.12.003>
- Mann, C. J. (2003). Observational research methods. Research design II: cohort, cross sectional, and case-control studies. *Emergency Medicine Journal*, 20(1), 54-60.
<http://dx.doi.org/10.1136/emj.20.1.54>
- Mayhorn, C. B., & Nyeste, P. G. (2012). Training users to counteract phishing. *Work*, 41. 3549-3552. <https://doi.org/10.3233/WOR-2012-1054-3549>

- McElwee, S., Murphy, G., & Shelton, P. (2018). Influencing outcomes and behaviors in simulated phishing exercises. In *SoutheastCon 2018*. 1-6.
<https://doi.org/10.1109/SECON.2018.8479109>
- Ndibwile, J. D., Luhanga, E. T., Fall, D., Miyamoto, D., Blanc, G., & Kadobayashi, Y. (2019). An empirical approach to phishing countermeasures through smart glasses and validation agents. *IEEE Access*, 7, 130758–130771. <https://doi.org/10.1109/access.2019.2940669>
- Nguyen, C., Jensen, M., & Day, E. (2023). Learning not to take the bait: a longitudinal examination of digital training methods and overlearning on phishing susceptibility. *European Journal of Information Systems*, 32(2), 238-262. <https://doi.org/10.1080/0960085X.2021.1931494>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 1-9.
<https://doi.org/10.1016/j.ijssu.2021.105906>
- Parsons, K., McCormac, A., Pattinson, M.R., Butavicius, M.A., & Jerram, C. (2013). Phishing for the truth: A scenario-based experiment of users' behavioural response to emails. *IFIP International Information Security Conference*. 405, 366-378. https://doi.org/10.1007/978-3-642-39218-4_27
- Parsons, K., McCormac, A., Pattinson, M., Butavicius, M., & Jerram, C. (2015). The design of phishing studies: Challenges for researchers. *Computers & Security*, 52, 194-206.
<https://doi.org/10.1016/j.cose.2015.02.008>
- Parsons, K., Butavicius, M., Delfabbro, P., & Lillie, M. (2019). Predicting susceptibility to social influence in phishing emails. *International Journal of Human-Computer Studies*, 128, 17-26.
<https://doi.org/10.1016/j.ijhcs.2019.02.007>
- Pattinson, M., Jerram, C., Parsons, K., McCormac, A., & Butavicius, M. (2012). Why do some people manage phishing e-mails better than others? *Information Management & Computer Security*, 20(1), 18-28. <https://doi.org/10.1108/09685221211219173>

- Phillips, J. K., Klein, G., & Sieck, W. R. (2004). Expertise in judgment and decision making: A case for training intuitive decision skills. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making*. 297–315. Blackwell Publishing.
<https://doi.org/10.1002/9780470752937.ch15>
- ProofPoint (2020). *2020 State of the phish: An in-depth exploration of user awareness, vulnerability and resilience*. <https://www.proofpoint.com/au/resources/threat-reports/state-of-phish>
- ProofPoint (2022). *2022 The definitive email cybersecurity strategy guide: A people-centric approach to stopping ransomware, malware attacks, phishing and email fraud*.
<https://www.proofpoint.com/au/resources/e-books/definitive-email-security-strategy-guide>
- ProofPoint (2023). *2023 State of the phish: An in-depth exploration of user awareness, vulnerability and resilience*. <https://www.proofpoint.com/au/resources/threat-reports/state-of-phish>
- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, & Cybernetics, SMC-13*(3), 257–266. <https://doi.org/10.1109/TSMC.1983.6313160>.
- Reeves, A., Calic, D., & Delfabbro, P. (2021a). “Get a red-hot poker and open up my eyes, it's so boring” 1: Employee perceptions of cybersecurity training. *Computers & Security, 106*, 1-13. 102281. <https://doi.org/10.1016/j.cose.2021.102281>
- Reeves, A., Delfabbro, P., & Calic, D. (2021b). Encouraging employee engagement with cybersecurity: How to tackle cyber fatigue. *SAGE Open, 11*(1), 1-18.
<https://doi.org/10.1177/21582440211000049>
- Reinheimer, B., Aldag, L., Mayer, P., Mossano, M., Duezguen, R., Lofthouse, B., ... & Volkamer, M. (2020). An investigation of phishing awareness and education over time: When and how to best remind users. In *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 259-284.

- Robinson, K. A., Akinyede, O., Dutta, T., Sawin, V. I., Li, T., Spencer, M. R., Turkelson, C. M., & Weston, C. (2013). *Framework for Determining Research Gaps During Systematic Review: Evaluation*. Agency for Healthcare Research and Quality (US).
- Roepke, R., Drury, V., Meyer, U., & Schroeder, U. (2022). Better the phish you know: Evaluating personalization in anti-phishing learning games. In *Proceedings of the 14th International Conference on Computer Supported Education (CSEDU 2022)*. 458-466.
<https://doi.org/10.5220/0011042100003182>
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52(1), 471-499. <https://doi.org/10.1146/annurev.psych.52.1.471>
- Sarno, D. M., McPherson, R., & Neider, M. B. (2022). Is the key to phishing training persistence? Developing a novel persistent intervention. *Journal of Experimental Psychology: Applied*, 28(1), 85-99. <https://doi.org/10.1037/xap0000410>
- Schoeb, G., Lafrenière-Carrier, B., Lauzier, M., & Courcy, F. (2021). Measuring transfer of training: Review and implications for future research. *Canadian Journal of Administrative Sciences*, 38(1), 17-28. <https://doi.org/10.1002/cjas.1577>
- Sharevski, F., & Jachim, P. (2022). "Alexa, what's a phishing email?": Training users to spot phishing emails using a voice assistant. *EURASIP Journal on Information Security*, 2022(7), 1-13. <https://doi.org/10.1186/s13635-022-00133-w>
- 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. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 373-382.
<https://doi.org/10.1145/1753326.1753383>
- Silic, M., & Lowry, P. B. (2020). Using design-science based gamification to improve organizational security training and compliance. *Journal of Management Information Systems*, 37(1), 129-161. <https://doi.org/10.1080/07421222.2019.1705512>

- Singh, K., Aggarwal, P., Rajivan, P., & Gonzalez, C. (2023). Cognitive elements of learning and discriminability in anti-phishing training. *Computers & Security*, 127, 103105. 1-15.
<https://doi.org/10.1016/j.cose.2023.103105>
- Stockhardt, S., Reinheimer, B., Volkamer, M., Mayer, P., Kunz, A., Rack, P., & Lehmann, D. (2016). Teaching phishing-security: which way is best? In *IFIP Advances in Information and Communication Technology* 135–149. https://doi.org/10.1007/978-3-319-33630-5_10
- Sturman, D., Valenzuela, C., Plate, O., Tanvir, T., Auton, J. C., Bayl-Smith, P., & Wiggins, M. W. (2023). The role of cue utilization in the detection of phishing emails. *Applied Ergonomics*, 106, 1-13. <https://doi.org/10.1016/j.apergo.2022.103887>
- Sumner, A., Yuan, X., Anwar, M., & McBride, M. (2022). Examining factors impacting the effectiveness of anti-phishing trainings. *Journal of Computer Information Systems*, 62(5), 975-997. <https://doi.org/10.1080/08874417.2021.1955638>
- Sutter, T., Bozkir, A. S., Gehring, B., & Berlich, P. (2022). Avoiding the hook: influential factors of phishing awareness training on click-rates and a data-driven approach to predict email difficulty perception. *IEEE Access*, 10, 100540-100565.
<https://doi.org/10.1109/access.2022.3207272>
- Telstra Corporation. (2018). *Telstra Security Report 2018*.
https://insight.telstra.com.au/content/dam/insights/pdfs/Telstra_Security_Report_2018_PDF_FINAL.PDF
- Tioh, J. N., Mina, M., & Jacobson, D. W. (2017). Cyber security training a survey of serious games in cyber security. In *2017 IEEE Frontiers in Education Conference (FIE)*. 1-5.
<https://doi.org/10.1109/FIE.2017.8190712>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, 169(7), 467-473. <https://doi.org/10.7326/M18-0850>

- Tschakert, K. F., & Ngamsuriyaroj, S. (2019). Effectiveness of and user preferences for security awareness training methodologies. *Heliyon*, 5(6), 1-10.
<https://doi.org/10.1016/j.heliyon.2019.e02010>
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124-1131. <https://doi.org/10.1016/B978-0-12-214850-7.50008-5>
- Valentine, J.A., 2006. Enhancing the employee security awareness model. *Computer Fraud & Security*, 6, 17–19. [https://doi.org/10.1016/S1361-3723\(06\)70370-0](https://doi.org/10.1016/S1361-3723(06)70370-0)
- Van Steen, T., Norris, E., Atha, K., & Joinson, A. (2020). What (if any) behaviour change techniques do government-led cybersecurity awareness campaigns use? *Journal of Cybersecurity*, 6(1), 1-8. <https://doi.org/10.1093/cybsec/tyaa019>
- Vayansky, I., & Kumar, S. (2018). Phishing—challenges and solutions. *Computer Fraud & Security*, 2018(1), 15-20. [https://doi.org/10.1016/S1361-3723\(18\)30007-1](https://doi.org/10.1016/S1361-3723(18)30007-1)
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Family Medicine*, 37(5), 360-363.
- Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, 51(3), 576-586.
<https://doi.org/10.1016/j.dss.2011.03.002>
- Vishwanath, A., Harrison, B., & Ng, Y. J. (2018). Suspicion, cognition, and automaticity model of phishing susceptibility. *Communication Research*, 45(8), 1146-1166.
<https://doi.org/10.1177/0093650215627>
- Volkamer, M., Renaud, K., Reinheimer, B., Rack, P., Ghiglieri, M., Mayer, P., Kunz, A., & Gerber, N. (2018). Developing and evaluating a five minute phishing awareness video. In *Trust, Privacy and Security in Digital Business*. 119–134. Springer International Publishing.
https://doi.org/10.1007/978-3-319-98385-1_9

- Wash, R., & Cooper, M. M. (2018, April). Who provides phishing training? facts, stories, and people like me. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 1-12. <https://doi.org/10.1145/3173574.3174066>
- Weaver, B. W., Braly, A. M., & Lane, D. M. (2021). Training users to identify phishing emails. *Journal of Educational Computing Research*, 59(6), 1169-1183. <https://doi.org/10.1177/0735633121992516>
- Wen, Z. A., Lin, Z., Chen, R., & Andersen, E. (2019, May). What.Hack: engaging anti-phishing training through a role-playing phishing simulation game. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1-12. <https://doi.org/10.1145/3290605.3300338>
- Wiggins, M., & O'Hare, D. (2003). Weatherwise: Evaluation of a cue-based training approach for the recognition of deteriorating weather conditions during flight. *Human Factors*, 45(2), 337-345. <https://doi.org/10.1518/hfes.45.2.337.27246>
- Williams, E. J., Hinds, J., & Joinson, A. N. (2018). Exploring susceptibility to phishing in the workplace. *International Journal of Human-Computer Studies*, 120, 1-13. <https://doi.org/10.1016/j.ijhcs.2018.06.004>
- Wolfe, J. M., Brunelli, D. N., Rubinstein, J., & Horowitz, T. S. (2013). Prevalence effects in newly trained airport checkpoint screeners: Trained observers miss rare targets, too. *Journal of Vision*, 13(3), 1-9. <https://doi.org/10.1167/13.3.33>
- Yang, W., Xiong, A., Chen, J., Proctor, R. W., & Li, N. (2017). Use of phishing training to improve security warning compliance: Evidence from a field experiment. In *Proceedings of the Hot Topics in Science of Security: Symposium and Bootcamp*. 52-61. <https://doi.org/10.1145/3055305.3055310>
- Yeoh, W., Huang, H., Lee, W.-S., Al Jafari, F., & Mansson, R. (2022). Simulated phishing attack and embedded training campaign. *Journal of Computer Information Systems*, 62(4), 802–821. <https://doi.org/10.1080/08874417.2021.1919941>

Zielinska, O. A., Tembe, R., Hong, K. W., Ge, X., Murphy-Hill, E., & Mayhorn, C. B. (2014). One phish, two phish, how to avoid the internet phish: Analysis of training strategies to detect phishing emails. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. 58(1), 1466-1470. <https://doi.org/10.1177/1541931214581306>

Appendix A

Journal Formatting Requirements

For ease of accessibility, the following paragraphs are direct quotes extracted from the Computers & Security Guide for Authors. If further detail is required, the document can be accessed via:

<https://www.elsevier.com/journals/computers-and-security/0167-4048/guide-for-authors#:~:text=All%20contributions%20should%20be%20in,speaker%20to%20avoid%20grammatical%20errors>.

“There are no strict formatting requirements, but all manuscripts must contain the essential elements needed to convey your manuscript, for example Abstract, Keywords, Introduction, Materials and Methods, Results, Conclusions, Artwork and Tables with Captions. If your article includes any Videos and/or other Supplementary material, this should be included in your initial submission for peer review purposes. Divide the article into clearly defined sections.”

“There are no strict requirements on reference formatting at submission. References can be in any style or format as long as the style is consistent. Where applicable, author(s) name(s), journal title/book title, chapter title/article title, year of publication, volume number/book chapter and the article number or pagination must be present. Use of DOI is highly encouraged.”

Appendix B

PRISMA-ScR Checklist

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
TITLE			
Title	1	Identify the report as a scoping review.	i
ABSTRACT			
Structured summary	2	Provide a structured summary that includes (as applicable): background, objectives, eligibility criteria, sources of evidence, charting methods, results, and conclusions that relate to the review questions and objectives.	vii
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known. Explain why the review questions/objectives lend themselves to a scoping review approach.	1-5
Objectives	4	Provide an explicit statement of the questions and objectives being addressed with reference to their key elements (e.g., population or participants, concepts, and context) or other relevant key elements used to conceptualize the review questions and/or objectives.	5
METHODS			
Protocol and registration	5	Indicate whether a review protocol exists; state if and where it can be accessed (e.g., a Web address); and if available, provide registration information, including the registration number.	6
Eligibility criteria	6	Specify characteristics of the sources of evidence used as eligibility criteria (e.g., years considered, language, and publication status), and provide a rationale.	7
Information sources*	7	Describe all information sources in the search (e.g., databases with dates of coverage and contact with authors to identify additional sources), as well as the date the most recent search was executed.	6-7
Search	8	Present the full electronic search strategy for at least 1 database, including any limits used, such that it could be repeated.	6
Selection of sources of evidence†	9	State the process for selecting sources of evidence (i.e., screening and eligibility) included in the scoping review.	7-8
Data charting process‡	10	Describe the methods of charting data from the included sources of evidence (e.g., calibrated forms or forms that have been tested by the team before their use, and whether data charting was done independently or in duplicate) and any processes for obtaining and confirming data from investigators.	9
Data items	11	List and define all variables for which data were sought and any assumptions and simplifications made.	9
Critical appraisal of individual sources of evidence§	12	If done, provide a rationale for conducting a critical appraisal of included sources of evidence; describe the methods used and how this information was used in any data synthesis (if appropriate).	9
Synthesis of results	13	Describe the methods of handling and summarizing the data that were charted.	9

SECTION	ITEM	PRISMA-ScR CHECKLIST ITEM	REPORTED ON PAGE #
RESULTS			
Selection of sources of evidence	14	Give numbers of sources of evidence screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally using a flow diagram.	8
Characteristics of sources of evidence	15	For each source of evidence, present characteristics for which data were charted and provide the citations.	9-10
Critical appraisal within sources of evidence	16	If done, present data on critical appraisal of included sources of evidence (see item 12).	11-18
Results of individual sources of evidence	17	For each included source of evidence, present the relevant data that were charted that relate to the review questions and objectives.	11-18
Synthesis of results	18	Summarize and/or present the charting results as they relate to the review questions and objectives.	19-30
DISCUSSION			
Summary of evidence	19	Summarize the main results (including an overview of concepts, themes, and types of evidence available), link to the review questions and objectives, and consider the relevance to key groups.	30-37
Limitations	20	Discuss the limitations of the scoping review process.	37-38
Conclusions	21	Provide a general interpretation of the results with respect to the review questions and objectives, as well as potential implications and/or next steps.	38-39
FUNDING			
Funding	22	Describe sources of funding for the included sources of evidence, as well as sources of funding for the scoping review. Describe the role of the funders of the scoping review.	N/A

JBI = Joanna Briggs Institute; PRISMA-ScR = Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews.

* Where *sources of evidence* (see second footnote) are compiled from, such as bibliographic databases, social media platforms, and Web sites.

† A more inclusive/heterogeneous term used to account for the different types of evidence or data sources (e.g., quantitative and/or qualitative research, expert opinion, and policy documents) that may be eligible in a scoping review as opposed to only studies. This is not to be confused with *information sources* (see first footnote).

‡ The frameworks by Arksey and O'Malley (6) and Levac and colleagues (7) and the JBI guidance (4, 5) refer to the process of data extraction in a scoping review as data charting.

§ The process of systematically examining research evidence to assess its validity, results, and relevance before using it to inform a decision. This term is used for items 12 and 19 instead of "risk of bias" (which is more applicable to systematic reviews of interventions) to include and acknowledge the various sources of evidence that may be used in a scoping review (e.g., quantitative and/or qualitative research, expert opinion, and policy document).

From: Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169:467–473. doi: 10.7326/M18-0850.