THE UNIVERSITY of ADELAIDE

User-Centric Design, Implementation and Evaluation Support for Phishing Interventions

Orvila Sarker

School of Computer and Mathematical Sciences
Faculty of Sciences, Engineering and Technology

Principal Supervisor: Dr. Asangi Jayatilaka
Co-Supervisors: Dr. Chelsea Liu, Dr. Sherif Haggag

A thesis submitted for the degree of
DOCTOR OF PHILOSOPHY
The University of Adelaide

December 25, 2023

Contents

List of Figures

List of Tables

Abstract

Phishing education, training, and awareness interventions are crucial to safeguarding organizations against malicious phishing attacks. However, the effectiveness of phishing interventions can be impeded by a lack of consideration of end-users' requirements and preferences by the practitioners during the design, implementation, and evaluation of these interventions. Such deficiency can result in user dissatisfaction, ineffectiveness of phishing interventions, and susceptibility of the intended end-users to phishing attacks. Failures to incorporate socio-technical issues during the design, implementation, or evaluation of phishing interventions often result from the unavailability of structured, personalized, and reliable guidance for the developers and practitioners of these interventions. Furthermore, the practical implementation of these guidelines is not without obstacles. To date, no study has provided personalized guidelines to support practitioners in addressing the challenges encountered in the design, implementation, or evaluation of phishing interventions. Additionally, no study has assessed the impediments associated with implementing academic guidelines within real-world settings.

The goal of this thesis is to address the current lack of resources and personalized guidelines for the design, implementation, and evaluation of anti-phishing interventions. This thesis systematically groups the scattered recommendations from the academic and grey literature to provide a list of organized and easily accessible recommendations for practitioners. To achieve the aforementioned goal, this research (i) systematically identified 20 challenges and 23 critical success factors within the design, implementation, and evaluation of phishing interventions from 53 academic and 16 grey literature studies; (ii) reports 22 socio-technical factors at the individual, technical, and organizational levels, that affected the effectiveness of anti-phishing interventions and require to tailor the phishing interventions; (iii) presents 41 guidelines personalized across 4 practitioner groups and 14 intervention types to address the identified challenges and socio-technical factors to improve the outcome of phishing interventions; (iv) provides an overview of the current anti-phishing defense mechanisms deployed in the organizations; (v) identifies 8 challenges faced by the practitioners in the design, implementation and evaluation of phishing interventions in real-world settings; (vi) investigates practitioners' perspectives on the devised guidelines to understand these guidelines' usefulness and applicability in practice; (vii) extracts features for an envisioned tool for practitioners preferences to easily access the reported guidelines. This thesis can be a valuable resource for the design, implementation, and evaluation of phishing interventions. The overarching goal is to augment the efficacy and success rates of these endeavors, thereby fortifying organizational defenses against sophisticated phishing attacks.

Declaration of Authorship

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that the copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis 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 University to restrict access for a period of time.

Orvila Sarker

December 2023

Acknowledgements

First, I express my sincere gratitude and profound appreciation to my principal supervisor, Dr. Asangi Jayatilaka, for providing me with constructive feedback on my academic progress and critically reviewing my work. I appreciate her commendable commitment to keeping me on track with my strict PhD timeline after the change of my research direction. Undoubtedly, Asangi is one of the most exceptional collaborators with whom I have had the privilege to work.

I am immensely thankful for the encouragement and motivation provided by my amazing and supportive co-supervisors Dr. Chelsea Liu, and Dr. Sherif Haggag. Chelsea has comprehensively reviewed all of my submissions within a constrained time frame. Her eloquence and outstanding presentation skills have profoundly inspired my professional development. I appreciate her valuable suggestions on how to effectively communicate research to a broader audience in a simplified manner. Sherif has always shown a willingness to support me. He has always expressed huge appreciation for my minor endeavors. In-person conversations with an approachable and positive person like him have always been rejuvenating.

My heartfelt thanks to my former PhD supervisors Prof. Hong Shen and Prof. M. Ali Babar for the valuable discussions on the research problems and for their professional advice. Prof. Ali's support has been instrumental in enabling me to stay focused and committed to the pursuit of my PhD.

My deep appreciation to The University of Adelaide and the Cyber Security Cooperative Research Centre (CSCRC) for funding my research. I also thank the People's Republic of Bangladesh for providing me with the opportunity to pursue higher education abroad.

I am thankful to the former and current members of the Centre for Research on Engineering Software Technologies (CREST) for their support. Special thanks to Chadni and Zahra for the valuable suggestions, and encouragement. Thanks to Roshan and Willam, for their consistent support during difficult times, and for always listening to my professional challenges.

Thanks to my friends Afifah, Manar, and Adrian for bringing out the best version of me and being there for me through thick and thin.

Finally, I express profound gratitude to my family for their unwavering support and selfless sacrifices. I thank my father, whose steadfast belief in my abilities since my childhood has been a continual source of encouragement. His lessons on staying calm and patient during challenging times have been crucial in helping me persist through the difficulties of my PhD journey. My heartfelt appreciation goes to my mother for her incredible affection and unwavering concern. Thanks to my sister, who has been my biggest inspiration and cheerleader since childhood. Her motivation, continuous presence, and reassurance have been invaluable. The accomplishments I have achieved would not have been possible without her. I express gratitude to my brother for his tireless efforts to resolve my problems and protect me from difficulties. Thanks to my brother-in-law for his understanding and support.

Dedicated to
My beloved father, thank you for your unconditional love, and sacrifices for your children. I will hold you in my heart until I can hold you in paradise.
Also, to my dear mother, thank you for your patience, affection, and kindness.

Chapter 1

Introduction

1.1 Background

Phishing is a form of cyber attack where attackers steal a user's personal information (e.g., online banking details, passwords, credit cards) through deception, such as by sending fraudulent emails or text messages). A recent data breach investigation report has revealed that 67% of successful cyber attacks are the result of human negligence, such as in the case of phishing [1]. In 2019, organizations lost more than US$1 billion to email phishing attacks [2]. In a recent survey of 7500 employees and 1050 cyber security professionals across 15 countries, it is observed that a total of 84% of surveyed organizations had experienced at least one successful phishing attack in 2022 [3]. According to recent APWG reports, 2023 was the worst year for phishing with a total of 1,077,501 phishing attacks being observed causing a loss of $56,195 per attempt [4]. Organizations typically use automated anti-phishing solutions to detect phishing attempts and safeguard their employees. Unfortunately, automated solutions (e.g., [5], [6], [7]), which are often built on probabilistic algorithms, suffer from false positives and false negatives (for instance PhishNot [5] generates 2.18% false positives and 3.22% false negatives in identifying phishing attempts). This highlights the importance of adopting human-centric phishing defense mechanisms to mitigate phishing risks [8]–[11].

One primary factor contributing to organizations' vulnerability to phishing attacks is the lack of awareness and knowledge regarding phishing attacks, particularly tactics, on the part of their employees [12]. As phishing primarily targets end-users, research highlights the need to provide more attention to user-centric anti-phishing defense in addition to the utilization of automated solutions (e.g., automated email filters, regular software system security updates, and malware scanning) [13] [12]. Moreover, user phishing education, training, and awareness programs are also considered the best defense against phishing attacks [14], [15].

A phishing intervention[1] is any anti-phishing mechanism, software application, tool, or framework designed to assist users in mitigating the impact of phishing attacks necessitating user involvement [21]. Phishing interventions can be broadly categorized as phishing education, training, and awareness as defined in Table 1.1.

[1]In this thesis, the terms phishing intervention and anti-phishing intervention are used interchangeably

TABLE 1.1. Main categories of phishing interventions in the scope of this thesis [16] [17]

Topic	Definition
Phishing education	Educational interventions are designed to cultivate user knowledge and proficiency in identifying phishing attempts, augmenting user comprehension, and empowering users to acquire insights into the phenomenon of phishing. Example: educational games [18].
Phishing training	The objective of training interventions is to furnish users with interactive and pragmatic experiences on safeguarding themselves against phishing attempts. Example: phishing simulation and embedded training [19].
Phishing awareness	Awareness interventions are employed as disruptions within users' typical workflow with the aim of issuing alerts and enhancing awareness regarding potential occurrences of phishing attacks. Often, these interventions offer users various design options for recognizing and identifying phishing attacks, such as the inclusion of custom icons, trust logos, and sender highlighting. Example: browser SSL warning [20].

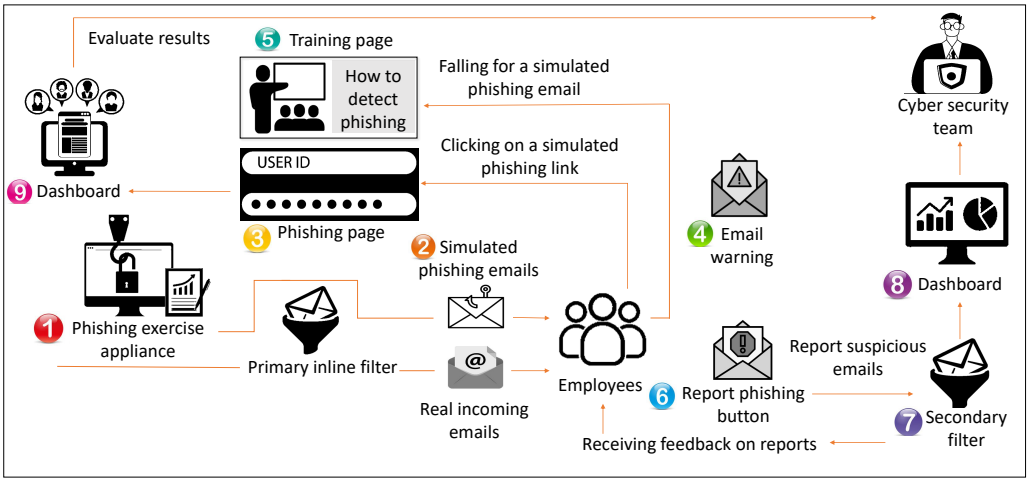

FIGURE 1.1. Overview of the design, implementation, and evaluation phase of a typical phishing training intervention

Figure 1.1 demonstrates different steps involved in the design, implementation, and evaluation of phishing simulation and embedded training intervention adopted from [22]. During the ***design*** stage (step 1 in Figure 1.1), the organization's security team i) defines the objectives of the phishing simulation and training intervention, ii) assess the current level of awareness and susceptibility of employees to phishing, iii) define the scope, target employees, and frequency of the simulation required, and iv) design the training contents and simulation emails based on realistic scenarios. During the ***implementation*** stage (step 2 to step 7 in Figure 1.1), the simulation emails are sent to the target employees. Subsequently, employees who click on the simulated phishing link receive an email warning, alerting them to the potential phishing threat. Those employees who persist in clicking on the simulated links are redirected to a training page. This instructional page contains information explaining the nature of the threat

and the key indicators of phishing they missed. Employees who correctly identify the simulated emails and report simulated phishing emails receive immediate acknowledgment from the system, thus reinforcing their vigilance against such threats. During the ***evaluation*** phase (step 8 and step 9 in Figure 1.1), the organization's security team establish key performance indicators (KPIs) to measure the effectiveness of the phishing simulation and training, analyze the results and feedback from each simulation to identify the areas for improvement, generate reports on user performance, and share insights with relevant stakeholders to demonstrate the intervention's impact on enhancing phishing awareness and resilience.

1.2 Research Motivation

Although many anti-phishing interventions are available, ranging from anti-phishing game [23], browser phishing warning [10], to phishing simulation and embedded training [24], end-users still fall victim to phishing attacks. The outcome and success of phishing interventions vary substantially on their design, implementation, and evaluation [19], [21], [25], [26]. Research has highlighted the importance of effective design, implementation, and evaluation of phishing interventions to transform end-users from a potential source of vulnerability to the strongest line of defense [27], [28]. Several empirical investigations have demonstrated that personalized design, implementation, and evaluation of phishing interventions to the needs of individual users can be effective in assisting end users in identifying and mitigating phishing attacks [9], [25], [29]–[34]. One-size-fits-all interventions would result in usability issues or human-centric weaknesses leading to end-users' susceptibility to phishing attacks [35]–[37].

Research has shown that end-users' preferences and socio-technical[2] factors can significantly impact the success of phishing interventions. For example, the perceived origin of phishing training material has a large impact on end-users' security outcomes: narrative-based training methods are more effective when told by a peer, and facts-and-advice-based training is more effective when presented by a security expert [39]. The importance of considering the target end-users' demographic for phishing training has also been emphasized: college students learn better from facts-based training from peers [40]. Casual gamers prefer simplified phishing educational games whereas serious gamers prefer congruent narratives [41].

As the needs, requirements, and preferences of end-users play an important role in determining the effectiveness of phishing intervention, practitioners[3] involved in the design, implementation, and evaluation of phishing interventions should consider end-users cognitive limitations and decision-making processes

[2] The term socio-technical can be described as "where the social and technical aspects are interwoven in a way that studying one without due consideration of the other makes for an incomplete investigation and understanding. Requirements elicitation and user-centered design are examples of socio-technical phenomena, where a complete research investigation needs to consider the full socio-technical context"[38].

[3] In this thesis the term "practitioner" refers to the people who are the designer/implementer/evaluator of anti-phishing software, tools or interventions. Examples of designers are browser warning developers, anti-phishing tool developers, email client phishing warning designers, phishing training program designers, and anti-phishing game developers. An example of an implementer or evaluator of phishing tools and training programs can be an information security officer of an organization. We also use the term practitioner to refer to the C-suite employees who are involved in the cyber security related decision-making of an organization or multiple organizations, for example, Chief Executive Officer (CEO), Manager, Cyber Security Experts.

in the design, implementation, and evaluation of phishing interventions [42]. Nevertheless, studies have demonstrated that practitioners often neglect the end-user preferences, mental states, and cognitive requirements in the design, implementation, and evaluation of anti-phishing interventions [23], [34]–[36], [43]–[54].

Several empirical investigations have revealed the failures of the practitioners in considering the needs and challenges faced by end-users in the decision-making of the design, implementation, and evaluation of phishing interventions [23], [34]–[36], [43]–[54]. For example, developers design and deploy passive phishing indicators that do not require user interruption and action. As a result, users often miss these indicators and fall victim to phishing [45]. Many email providers do not use Simple Mail Transfer Protocol (SMTP) authentication mechanisms. Attackers use this as an opportunity to send phishing emails from spoofed email addresses [43]. Despite the severe consequences of phishing attacks, the majority of cyber security training programs are designed without detailed coverage of phishing-related knowledge [44]. Designers frequently eliminate crucial phishing indicators within mobile operating systems and browsers to accommodate additional content within the limited display dimensions of smartphones and tablet PCs [55]. Digital service providers, such as banks, provide abstract and conflicting anti-phishing recommendations, which are not helpful for end users [46]. Security practitioners often design phishing training with unrealistic or irrelevant content and do not follow any formal procedures [35], [47]–[49]. Intervention designers design interventions with complicated interfaces that require user special knowledge to install and use [23], [34], [50], [51].

The aforementioned examples underscore the necessity for a systematic set of guidelines for the practitioners, facilitating their comprehension of the diverse requirements and challenges encountered by end-users. Unfortunately, to the best of our knowledge, there are no structured and personalized guidelines available for the practitioners involved in the design, implementation, and evaluation of phishing interventions. Some existing resources are intended only for developers and are not specific to phishing interventions. For instance, Lujo et al. [56] and Lynsay et al. [57] reported guidelines for browser security warning design (fall under a particular class of phishing intervention), and Mirium et al. [58] provided guidelines for the development of cyber security games (not directly related to phishing). Also, the target user group of these guidelines is primarily developers, with limited to no relevance to other security practitioners.

Research has recommended that the guidelines proposed in the literature should be evaluated to investigate and understand their applicability and usefulness in practice [59]. Guidelines or recommendations need empirical evaluation to capture implementation details necessary for practitioners to feel confident to adopt them [60], [61]. Understanding the challenges in translating academic research into practical application is important, as the applicability and utility of academic results and findings may not always align seamlessly with practice [48], [60], [62]. In the context of phishing education, training, and awareness,

understanding the current challenges, and identifying obstacles to implementing academic guidelines is crucial for several reasons. Research conducted in real industry settings underscores the challenges entailed in the implementation of anti-phishing interventions, for example, deploying phishing education and training programs in practice [63], [64], and deploying anti-phishing browser plug-ins in different browser platforms [65]. Real-world implementation challenges arise mainly due to distinct design patterns of browser platforms and devices among vendors and the unique settings and requirements of different organizations. For example, Small and medium enterprises (SMEs) may exhibit distinct requisites owing to resource constraints, characterized by a reduced number of security officers and specialists available to scrutinize the efficacy of anti-phishing defense mechanisms within their organizational frameworks [66]. Also, understanding the challenges in practice would be beneficial as guidelines devised from academic studies that sometimes lack clear external validity, as most of them considered small sample sizes [67]–[69], little diversity [68], [69], or role-playing scenarios [19], [68]. Moreover, sometimes academic findings contradict the findings identified in real industry settings [22], [63]. This highlights the importance of investigating the challenges in practical organizational settings and evaluating the guidelines in practice to understand practitioners' needs.

Discovering relevant guidelines can be an overwhelming task for practitioners, particularly when these guidelines lack organization and personalization. Several justifications underlie this difficulty. Diverse types of interventions, such as educational games, embedded training, and browser warnings, pose distinct challenges. To illustrate, challenges associated with end-users' retention of phishing knowledge and the assessment of such knowledge are mainly germane to interventions focusing on phishing education and training (e.g., [70], [69]). Again, certain challenges are contingent upon specific stages of intervention implementation; for instance, challenges related to the deployment of anti-phishing technology and automation, as discussed in the literature ([47], [65]), commonly relate to the implementation phases of interventions and challenges related to UI interfaces are relevant to design phases of the interventions (e.g., [35], [36], [53]). In light of these examples, we posit that guidelines can be effectively communicated to practitioners through a personalized tool. A tool can support them in finding guidelines specific to intervention types, stages, or related challenges to the intervention.

Problem Statement: Current one-size-fits-all phishing interventions remain ineffective, which is attributable to the practitioners overlooking the needs and challenges end-users face throughout the stages of design, implementation, and evaluation. This inadequacy manifests in user dissatisfaction, diminished efficacy, and heightened vulnerability of end-users to phishing attacks. The omission of socio-technical considerations in interventions arises from two main reasons: firstly, the unavailability of structured and personalized guidelines for the practitioners, including evaluation of the usefulness, applicability, and impediments to implementing these guidelines in practical organizational settings; and secondly, a lack of comprehension and understanding regarding the challenges and requirements of the practitioners in the design, implementation, and evaluation of anti-phishing interventions in practice.

1.3 Objectives and Research Questions

In the design, implementation, and evaluation of anti-phishing interventions, there is a tendency among practitioners to overlook the needs and preferences of end-users. The omission of consideration for socio-technical challenges and issues in the design, implementation, and evaluation phases renders these interventions ineffective, consequently exposing the targeted end-users to vulnerabilities associated with phishing attacks. This disregard for socio-technical challenges and issues can be attributed to the absence of structured and personalized guidelines for the practitioners.

This thesis aims to provide practitioners with guidance and support in facilitating the incorporation of socio-technical challenges and issues into phishing interventions. Additionally, it seeks to understand the practical challenges encountered by these professionals and elucidate how the provided guidelines can be rendered beneficial, applicable, and easily accessible within their respective organizational contexts.

💬RQ1. What are the socio-technical challenges in designing, implementing, and evaluating phishing interventions reported in the literature?

To integrate the requirements and preferences of users into phishing interventions, practitioners need to possess a thorough comprehension of the socio-technical challenges inherent in the stages of design, implementation, and evaluation. Despite a notable demand for a comprehensive body of knowledge addressing the challenges associated with phishing interventions, there has been, to the best of our knowledge, no effort to systematically organize these challenges. Consequently, we formulate RQ1 to ascertain the constraints and obstacles throughout the stages of design, implementation, and evaluation.

💬RQ2. What are the critical success factors reported in the literature for designing, implementing, and evaluating phishing interventions?

Existing literature has offered recommendations and discussed critical success factors derived from empirical investigations and experiential insights for mitigating usability issues associated with phishing interventions. Unfortunately, these insights are scattered across diverse academic and grey literature studies, lacking systematic organization in an easily accessible format. To bridge this gap, RQ2 endeavors to discern potentially critical success factors that may furnish valuable insights for practitioners to enhance the efficacy of phishing interventions.

💬RQ3. What are the socio-technical factors that impact the effectiveness of phishing interventions in the design, implementation, and evaluation stages?

Having identified the challenges and critical success factors to address RQ1 and RQ2, our investigation delved into primary determinants associated with sub-optimal outcomes in anti-phishing interventions. Therefore, we investigated the main reasons for poor outcomes of existing phishing interventions, as well as areas of improvement suggested by the studies to achieve a better user experience. Through the synthesis of this accumulated information, RQ3 aims to identify individual, technical, and organizational factors that contribute to either enhancing or impeding the overall efficacy of anti-phishing interventions.

💬RQ4. What guidance can be provided to support the practitioners in addressing and incorporating the socio-technical challenges and factors in anti-phishing interventions?

Practitioners lack a systematic set of guidelines that facilitate their comprehension of the diverse requirements and challenges encountered by end-users when designing, implementing, and evaluating phishing interventions resulting in one-size-fits-all solutions. Therefore, the main motivation behind RQ4 is to devise structured and personalized guidelines to support the practitioners in tailoring the design, implementation, and evaluation of phishing interventions for them to be effective for diverse groups of end-users, their needs, and challenges.

💬RQ5. What are the challenges, requirements, and preferences of the practitioners in adopting the identified guidelines in practice?

The adoption and implementation of the devised guidelines from the literature need to be evaluated by the practitioners to understand their practical usefulness and applicability. To answer this research question, we formulated the following four sub-research questions:

💬RQ5.1. What are the current anti-phishing practices employed by organizations?

From the answers to RQ5.1, our primary objective is to comprehend the existing anti-phishing practices within organizations. This endeavor affords the opportunity to understand the unique requirements of each organization and the required modifications essential for adapting the devised guidelines to diverse contexts and organizational settings.

💬RQ5.2. What are the socio-technical challenges perceived by the practitioners in designing, implementing, or evaluating anti-phishing interventions?

The formulated guidelines must effectively contend with the intricacies associated with designing, implementing, or evaluating interventions in practice. To

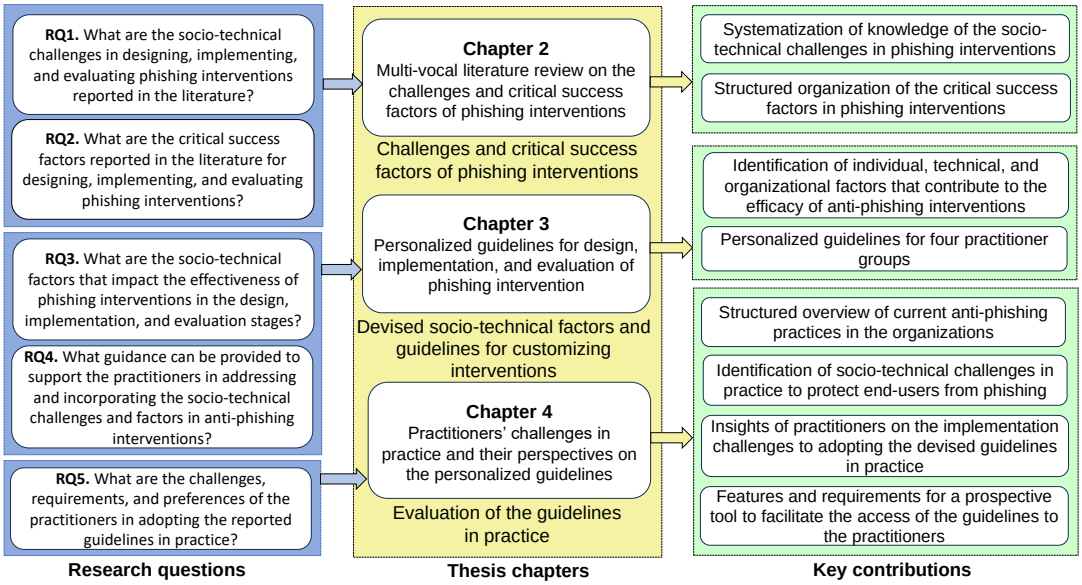

FIGURE 1.2. Overview of the thesis

attain this objective, RQ2 endeavors to scrutinize the challenges encountered by the practitioners, examining the extent to which challenges identified from the existing literature align with those encountered in real-world scenarios.

💬**RQ5.3. What are the perceptions of developers and practitioners regarding the usefulness and relevance of the reported guidelines?**

Guidelines developed from academic literature are required to be evaluated to understand their usefulness and applicability in practice. The objective of RQ5.3 is to understand the practitioners' perspectives on the strengths and weaknesses of the guidelines.

💬**RQ5.4. What features are desired by the practitioners for a tool that intends to facilitate them in accessing the guidelines?**

The devised guidelines need to be easily accessible and available in a format preferred by the practitioners to minimize the limitations of adopting them in practice. RQ5.4 aims to understand what features in the tool practitioners would prefer to access the guidelines.

1.4 Thesis Overview

Figure 1.2 displays the overview of the thesis (research questions under RQ5 are omitted for simplicity). The remainder of the thesis is organized as follows:

📖 ***Chapter 2.*** **Multi-vocal Literature Review on the Challenges and Critical Success Factors of Phishing Interventions**

We conduct a systematic multi-vocal literature review with 53 high-ranked academic studies and 16 high-quality grey studies, including several industry reports, industry case studies, and practitioner blogs to gather a comprehensive and structured overview of the challenges and critical success factors in the design, implementation, and evaluation of phishing education, training, and awareness interventions. The challenges discussed in this chapter provide the necessary justification to support the problem addressed in this thesis. More

specifically, it highlights the importance of tailored interventions and designing, implementing, and evaluating incorporating end-users' needs and preferences in the interventions for these to be effective for the diverse range of end-users.

Chapter 3. Personalized Guidelines for Design, Implementation, and Evaluation of Phishing Interventions

From the challenges and critical success factors discussed in Chapter 2, we systematically identified the factors that impact the effectiveness of the design, implementation, and evaluation of phishing interventions. We identify 22 socio-technical factors including end-users' demographics (e.g., age, educational qualification), mental states (e.g., security fatigue, pressure, optimism bias), and cognitive limitations (e.g., knowledge decay) which are currently ill-considered and require more attention from the practitioners to tailor the design, implementation, and evaluation phishing interventions. This chapter also presents a set of 41 guidelines personalized to four practitioner groups involved in the design, implementation, and evaluation of 14 types of phishing interventions. The guidelines are systematically devised from the critical success factors identified in Chapter 2. These guidelines are devised to support practitioners in addressing the socio-technical challenges identified in Chapter 2 and to incorporate socio-technical factors identified in this Chapter.

Chapter 4. Practitioners' Challenges in Practice and their Perspectives on the Personalized Guidelines

Providing a set of guidelines to the practitioners as discussed in Chapter 3 to facilitate tailored design, implementation, and evaluation of phishing interventions will not help reduce end-users' phishing risk until the practical implementation challenges of these guidelines are understood as well as practitioners challenges and requirements are investigated. Moreover, the guidelines devised in Chapter 3 aim to address the socio-technical challenges (discussed in Chapter 2) and socio-technical factors (discussed in Chapter 3) are mainly identified from the literature. Therefore it is important to evaluate the usefulness and applicability of these guidelines in practice. To achieve this, we conducted 18 semi-structured interviews with 18 practitioners of 18 organizations from 6 countries. This chapter provides a systematic overview of current anti-phishing practices in the industry to protect organizations from phishing attacks. Then this chapter discusses 8 ongoing challenges faced by the practitioners in the design, implementation, and evaluation of anti-phishing interventions. A comparison of these 8 challenges identified in practice with challenges identified in the literature (discussed in Chapter 2) is described in this chapter to understand to what extent the challenges in literature are aligned with the challenges in practice. Furthermore, this chapter includes practitioners' perspectives on the devised personalized guidelines (discussed in Chapter 3) and highlights the necessary areas to improve to make these guidelines useful and applicable to the industry practitioners. This chapter also collected recommendations provided by the practitioners for an intended tool to access the guidelines. This chapter sheds light on bridging the gap between academia and industry in anti-phishing efforts.

Chapter 5. Conclusion and Future Research Directions

This chapter summarizes the findings of this thesis. Also, five future research

directions are discussed in this chapter in addition to the recommendations provided in Chapter 2 and Chapter 4. In this chapter, the need to investigate the requirements of Small and Medium Enterprises (SMEs) for phishing defense based on the findings obtained in Chapter 4 is highlighted. This chapter also emphasizes that future studies can investigate the challenges identified in Chapter 2 and Chapter 4 rigorously in large-scale settings. Moreover, this chapter recommends the design of a tool for the automated generation of guidelines based on practitioners' input requirements and evaluating of the guidelines for a different set of settings such as with a different set of practitioners to understand barriers for practitioners' adoption and acceptance.

1.5 Thesis Contributions

❶ Systematization of knowledge of the socio-technical challenges in phishing interventions (**Chapter** 2).

We present the first-of-its-kind systematization of knowledge of the challenges including 8 design, 7 implementation, and 5 evaluation challenges in phishing interventions. Our analysis underscores several requirements that are currently missing or ill-considered in the design, implementation, and evaluation, for instance, the absence of active interruption in the phishing warnings and unsuitable warning placement. We highlight some implementation challenges, such as issues that arise due to browser platform dependency and distributed work settings in organizations. Our investigation reveals the usability issues incurred in phishing interventions due to a lack of evaluation of the diverse demographic features of end-users during the early prototype construction of the interventions. Overall, our endeavor aims to empower practitioners with a comprehension of the constraints associated with anti-phishing interventions.

❷ Structured organization of the critical success factors in phishing interventions (**Chapter** 2).

We make the first effort to systematically organize the scattered recommendations that act as determinants to potentially improve the effectiveness of phishing interventions. We identify 23 critical success factors in the design, implementation, and evaluation of phishing interventions including the design of diversified, up-to-date, and engaging training content, adoption of dynamic and self-adaptive training, and design of unified phishing indicators across different browser platforms and devices. Our investigation provides novel insights intended to augment the efficacy of anti-phishing initiatives within intricate real-world contexts.

❸ Identification of individual, technical, and organizational factors that contribute to the efficacy of anti-phishing interventions (**Chapter** 3).

We identify a novel set of 22 socio-technical factors including 15 human factors, 4 technical factors, and 3 organizational factors that are required to tailor the phishing interventions. This investigation reveals that customizing interventions requires considering several end-user demographics such as age, educational qualification, knowledge level, and organization position. Similarly, we observe users' cognitive constraints such as knowledge decay, security fatigue, and distraction need to be considered to make the interventions effective and useful to the end-users.

❹ Personalized guidelines for practitioners to address socio-technical challenges and to incorporate socio-technical factors in phishing interventions (**Chapter** 3).

We identify 41 guidelines for the design, implementation, and evaluation of phishing interventions. These guidelines are personalized in terms of 14 types of phishing interventions, 4 practitioner groups (designer/developer, information security team members of organizations, cyber security experts, and C-suite employees of organizations), 3 intervention stages (design, implementation, and evaluation), 19 challenges and 22 socio-technical factors. These guidelines are systematically compiled based on our identified 23 critical success factors. Our devised guidelines can be a useful resource to aid practitioners in improving the socio-technical challenges in the design, implementation, and evaluation of phishing interventions.

❺ Structured overview of current anti-phishing practices in the organizations (**Chapter** 4).

We provide a systematized view of anti-phishing practices employed in the organizations. This empirical observation includes the phishing training methods and contents used in the organizations, the frequency of training employed, the reminder and notification process used, the evaluation process performed for employees' knowledge assessment, and manual and automated phishing detection mechanisms deployed in the organizations and the actions performed in case of a real phishing attack. This knowledge can assist security experts and decision-makers in imposing the necessary changes required to protect organizations from phishing attacks.

❻ Identification of socio-technical challenges in practice to protect end-users from phishing (**Chapter** 4).

We derive an empirical understanding of the current challenges in the design, implementation, and evaluation and identify 8 challenges in practice. We compare and contrast these 8 challenges identified in practice with 19 challenges identified from the literature to understand the similarities and new findings. Our empirical analysis demonstrates real-world challenges such as obstacles in training content design, limitations of anti-phishing datasets, limitations of training materials, challenges to motivating employees to encourage

secure behavior, etc. Our investigation of these real-world challenges led us to understand how these challenges can be potentially addressed by our devised guidelines.

> ❼ Insights of practitioners on the implementation challenges to adopting the devised guidelines in practice (**Chapter** 4).

We report insights from industry practitioners to understand the strengths of our devised guidelines as well as the real-world implementation challenges associated with them. We systematically collect several recommendations on our guidelines and report the consequences of implementing our guidelines in different organizational contexts and settings. This empirical analysis enables us to understand the usefulness and applicability of our devised guidelines.

> ❽ Features and requirements for a prospective tool to facilitate the access of the guidelines to the practitioners (**Chapter** 4).

We summarize actionable features and requirements for an envisioned tool for the practitioners to access the guidelines. These features can guide researchers and tool developers to design tools incorporating features such as natural language processing-based search queries to generate guidelines and automatic updates of new guidelines.

1.6 Publications

1.6.1 Publication related to this thesis

The findings and contributions presented in this thesis are based on the following publications:

Orvila Sarker, Asangi Jayatilaka, Sherif Haggag, Chelsea Liu, and M. Ali Babar. "A Multi-vocal Literature Review on Challenges and Critical Success Factors of Phishing Education, Training and Awareness", *The Journal of Systems and Software*, 2024. [CORE ranking: rank A, Impact factor (2022): 3.5, SJR ranking: Q1]. (**Chapter** 2)

Orvila Sarker, Sherif Haggag, Asangi Jayatilaka, Chelsea Liu, "Personalized Guidelines for Design, Implementation and Evaluation of Anti-phishing Interventions", *17th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2023. [CORE ranking: rank A, acceptance rate 28%]. (**Chapter** 3)

Orvila Sarker, Asangi Jayatilaka, Sherif Haggag, Chelsea Liu, and M. Ali Babar "Understanding Practitioner's Challenges and Requirements in the Design, Implementation, and Evaluation of Anti-phishing Interventions", The Journal of Systems and Software (submitted). [CORE ranking: rank A, Impact factor (2022): 3.5, SJR ranking: Q1]. (**Chapter** 4)

1.6.2 Other publication

Aside from the aforementioned publications, it is noteworthy that an additional publication has been developed during the early stage of my PhD candidature which is not included in this thesis. It is important to highlight that during the first and second years of my PhD, I conducted research under the guidance of my former supervisors, Professor Hong Shen, and Professor M. Ali Babar, leading to the development of the following work. During this time, I have worked on developing security solutions for vehicular Adhoc Networks (VANETs). Subsequently, following the departure of my former principal supervisor, Professor Hong Shen, from his position at The University of Adelaide, I was directed to a new research topic namely phishing education, training, and awareness. The work produced on this new research topic in collaboration with my current supervisors is the main content of this thesis.

Orvila Sarker, Hong Shen, and M. Ali Babar "Reinforcement Learning Based Neighbour Selection for VANET with Adaptive Trust Management", *22nd IEEE International Conference on Trust, Security, and Privacy in Computing and Communications (TrustCom)*, 2023. [CORE ranking: rank B, acceptance rate 30%].

Statement of Authorship

Title of Paper	A multivocal literature review on challenges and critical success factors of phishing education, training, and awareness
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in a manuscript style
Publication Details	O. Sarker, A. Jayatilaka, S. Haggag, C. Liu, and M. A. Babar, "A multivocal literature review on challenges and critical success factors of phishing education, training and awareness," Journal of Systems and Software, vol. 208, p. 111 899, 2024.

Principal Author

Name of Principal Author (Candidate)	Orvila Sarker		
Contribution to the Paper	Performed analysis on all data, interpreted data, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	85%		
Signature		Date	20/12/2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Asangi Jayatilaka		
Contribution to the Paper	Supervised development of work, helped in developing methodology, data interpretation and edit the manuscript.		
Signature		Date	20/12/2023

Name of Co-Author	Sherif Haggag		
Contribution to the Paper	Supervised development of work, helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Please cut and paste additional co-author panels here as required.

Name of Co-Author	Chelsea Liu		
Contribution to the Paper	Supervised development of work, helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Name of Co-Author	M. Ali Babar		
Contribution to the Paper	Helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Chapter 2

Multi-vocal Literature Review on the Challenges and Critical Success Factors of Phishing Interventions

Related Publication: This chapter is based on our paper: *"A Multi-vocal Literature Review on Challenges and Critical Success Factors of Phishing Education, Training and Awareness", published in The Journal of Systems & Software.* (CORE ranking: rank A) [71]

Understanding the diverse challenges faced by the end-users of phishing education, training, and awareness (PETA) interventions and critical success factors documented in the literature to improve these interventions is one of the preliminary steps to making an effort to effective interventions. This chapter presents a comprehensive, structured view of the challenges and critical success factors of the design, implementation, and evaluation stages of phishing PETA. More specifically, a systematic Multi-vocal Literature Review (MLR) of 53 academic studies and 16 grey studies (including industry reports, industry case studies, and practitioners' blogs reflecting their experiences) from popular databases by following a well-known MLR guideline is reported in this chapter. We identified 20 socio-technical challenges and 23 critical success factors in the design, implementation, and evaluation stages of PETA. This chapter has enabled us to build a solid foundation highlighting the need for addressing socio-technical issues, the consequences of not considering/incorporating users' requirements in the interventions, and the importance of adopting personalized approaches in the design, implementation, and evaluation. Based on our experience conducting the study reported in this chapter, we discuss several open issues in PETA intervention such as the need for designing explainable anti-phishing systems and developing automated tools and platforms to conduct real-world phishing studies.

To the best of our knowledge, our review is the first to comprehensively analyze and synthesize the useful information and insights scattered across numerous studies in the academic and grey literature facilitating an in-depth understanding of the current challenges and critical success factors in the state-of-the-art and state-of-practice.

2.1 Introduction

Despite an increasing number of available anti-phishing interventions (e.g., anti-phishing game [23], browser phishing warning [10], anti-phishing training [24]), end-users still fall prey to phishing attacks. This is due to the reason that the success of anti-phishing interventions to educate and train the end-users can vary substantially depending on their design, implementation, and evaluation [19], [21], [25], [26]. Therefore, an effective design, implementation, and evaluation of anti-phishing intervention are required to enable organizations to turn their employees from a potential source of cyber security vulnerability into their strongest line of defense, by providing employees with the skills to identify and report phishing attacks [27], [28]. Consequently, to improve anti-phishing interventions, it would be valuable for designers and practitioners of anti-phishing interventions to be aware of the challenges and critical success factors associated with their design, implementation, and evaluation stages. Yet despite the rapidly growing body of academic research into anti-phishing interventions, to the best of our knowledge, there has been no review conducted aimed at organizing the body of knowledge on anti-phishing education, training, and awareness interventions to provide a comprehensive and in-depth synthesis of the existing evidence of the challenges and critical success factors that drive the effectiveness of anti-phishing interventions. Responding to this evident lack of investigation into an important topic, we conduct a Multi-Vocal Review (MLR) by systematically analyzing the peer-reviewed and grey literature on this topic.

Conducting an MLR offers several advantages: i) provides richer data and strong ecological validity in contrast to laboratory settings used in academic studies as indicated by Greene, Steves, Theofanos, *et al.* [72]; ii) allows a more comprehensive analysis for answering the relevant research questions by combining the state-of-the-art and the state-of-practice [73]; iii) given that anti-phishing interventions are inherently an industry-oriented practice, including the voice of practitioners ensure that practitioners' experience and industry viewpoints are not missed [74], [75]; iv) enables to canvas of abundant practical information from diverse document sources, such as (for example, phishing vendor manuals and guides to run phishing campaigns and evaluate phishing simulations [76]), which provide insights into real-world policies and practices.This study makes the following significant contributions:

- We provide an in-depth analysis of the challenges in PETA interventions to enable researchers and practitioners to gain a better understanding of the limitations of anti-phishing initiatives, which in turn helps improve their effectiveness in safeguarding organizations against future phishing attacks.
- We offer a comprehensive overview of critical factors in the design, implementation, and evaluation stages that determine the success of PETA interventions, providing researchers and practitioners with novel insights and guidance on how to enhance the success of anti-phishing initiatives in complex real-world contexts.
- We present a set of recommendations to guide researchers and practitioners, based on prior empirical evidence, to develop novel approaches to PETA to

TABLE 2.1. Comparison of our study with the related existing studies (- : topic not discussed, ✓ : topic discussed)

Contribution	Jampen, Gür, Sutter, *et al.* [77]	Franz, Zimmermann, Albrecht, *et al.* [21]	Our study
Challenges in PETA	-	-	✓
Critical factors in design, implementation, and evaluation phases of PETA	-	-	✓
Factors having impact on anti-phishing training	✓	-	✓
Phishing attack vector	-	✓	-
Time for PETA intervention reception	-	✓	-
Taxonomy of the existing solutions on user education, training, and awareness	-	✓	-
Future research directions	✓	-	✓
Study type	Survey	SLR	MLR

counter phishing attacks.

2.2 Related Work

PETA interventions (the terms education, training, and awareness are defined in Table 1.1) constitute an essential line of defense to mitigate phishing threats that bypass automated detection tools [25], [78]. The importance of anti-phishing interventions has attracted significant attention from researchers to conduct human-centric phishing studies. Nevertheless, despite the large and scattered body of evidence on PETA interventions, the prior research has yet to provide a comprehensive overview of the challenges and critical factors that determine the success of PETA interventions. Furthermore, the utmost importance of real-world industrial settings in investigating phishing was emphasized in several studies (e.g., Althobaiti, Jenkins, and Vaniea [47], Burda, Chotza, Allodi, *et al.* [79]). Replicating phishing studies to evaluate the effects of variations in different industrial settings was also recommended [79]. Despite being an industry-oriented domain, the inclusion of the practitioner perspective has been overlooked in previous reviews. Consequently, a comprehensive understanding of the prevailing challenges and essential factors for successful outcomes remains incomplete. This study aims to address this research gap by synthesizing a comprehensive body of knowledge derived from practical experiences in the industrial setting. This endeavor will encompass insights obtained from diverse sources of grey literature, effectively capturing the perspectives of practitioners.

A body of the existing work ([21], [77]) has attempted to provide an overview of the existing evidence, including the taxonomy of user-based PETA interventions, discussed various elements of the training materials. While the utilization of a taxonomy discussed by Franz, Zimmermann, Albrecht, *et al.* [21] aids in comprehending the existing interventions and their underlying mechanisms, as well as facilitating comparative analysis, it offers limited insight into the encountered challenges faced by these interventions. A comprehensive analysis of the prevailing challenges remains absent in the literature. A thorough comprehension of these challenges can greatly assist researchers and practitioners in devising more effective solutions to address or enhance the current issues

within these interventions. The purpose of our study is to bridge this gap by conducting a comprehensive investigation. Our investigation differs from the study conducted by Jampen, Gür, Sutter, *et al.* [77], which primarily delved into phishing training interventions, by examining a range of factors that influence the efficacy of phishing training. Our study, in contrast, concentrates on the challenges and pivotal elements that ascertain the effectiveness of PETA encompassing three distinct interventions: phishing education, training, and awareness.

The examination of user demographic information, methodology and evaluation techniques, and various human factors as presented by [80], [17], [81], and [16] has made a significant contribution to the existing body of knowledge. These scholarly works have provided valuable insights into the diverse characteristics and vulnerabilities of distinct user groups, thereby benefiting both researchers and practitioners. Nonetheless, for this knowledge to effectively inform the development of targeted and customized intervention strategies, a comprehensive overview of the specific challenges encountered by the end-users and the factors that contribute to successful interventions is necessary. It is the objective of our study to address this gap by presenting an inclusive analysis of the challenges faced and critical success factors associated with these interventions. The findings we report can aid intervention designers in tailoring educational materials, training programs, and communication strategies to effectively engage and educate diverse user groups. Table 2.1 summarizes the specific areas of contribution offered by our study in comparison to these prior studies.

2.3 Research Methodology

We follow the guidelines provided by Kitchenham and Charters [82] and Garousi, Felderer, and Mäntylä [83] to conduct our MLR.

2.3.1 Research questions

Our MLR aims to identify a comprehensive list of the challenges and critical success factors in the design, implementation, and evaluation stages of PETA. To achieve this we formulate the following research questions.

💬**RQ1. What are the socio-technical challenges in designing, implementing, and evaluating phishing interventions reported in the literature?**

Motivation: The motivation of this RQ is to understand the constraints and obstacles faced by researchers and practitioners in educating, training, and raising user awareness about phishing during the design, implementation, and evaluation stages.

💬**RQ2. What are the critical success factors reported in the literature for designing, implementing, and evaluating phishing interventions?**

Motivation: The primary motivation is to identify the potentially influential factors that can provide actionable insights to the researchers and practitioners to develop improved PETA interventions.

TABLE 2.2. Conceptualization of search strings (search strings are categorized based on the definition provided in Table 1.1)

Study	Focus areas	Search Strings	Source
Academic	Phishing Education Training Awareness	"phish*" *AND* ("educat*" *OR* "teach*" *OR* "learn*") *AND* "train*" *AND* ("aware*" *OR* "interven*" *OR* "nudge*" *OR* "warn*" *OR* "protect*" *OR* "security indicators" *OR* "alert*")	Scopus
Grey	Phishing education/ training/awareness	"phishing" *AND* ("education" *OR* "training" *OR* "awareness")	Google

2.3.2 Search strategy

Academic literature

To collect academic studies, we used Scopus[1] as our search database. The decision to employ Scopus search engine to identify the relevant primary studies was based on: i) the experiences reported by several other studies [84]–[87] justifying that Scopus indexes a large majority of the journals and conference papers in indexed by many other search engines such as IEEE Xplore, ACM Digital Library and SpringerLink ; ii) the fact that Scopus track a large number of journals and conferences in software engineering and computer science [84]–[87] that were the main target of this review; and iii) our pilot search results confirming the comprehensiveness of Scopus compared to other databases by verifying that no important studies were overlooked by solely relying on Scopus. With respect to the pilots conducted in other databases, we will explain the results of the pilot searches conducted in IEEE Xplore and ACM Digital Library databases as examples here. The search in IEEE Xplore yielded a total of 888 studies, after removing 52 duplicates. From this sample, 70 studies were randomly selected and cross-checked with Scopus search results. Of these, 64 studies were found in Scopus and the remaining 6 were not relevant to the study. Subsequently, 40 studies were identified from the ACM Digital Library, of which 25 were randomly selected and cross-checked with Scopus. Among these, 22 studies were found in Scopus, and the remaining 3 were not relevant to the study. Hence, based on the pilot search, we can conclude that no relevant studies were missed by using Scopus as the search engine. As a result, we are confident that using Scopus as the search engine could cover most of the relevant papers for this study. A similar observation was made by Kitchenham, Pretorius, Budgen, *et al.* [85] in their study. By comparing the results of Scopus with a manual search in their study, Kitchenham, Pretorius, Budgen, *et al.* [85] observed that Scopus covered all the relevant papers that used appropriate terminology.

[1] https://www.scopus.com/

TABLE 2.3. The inclusion and exclusion criteria

Inclusion criteria	**Exclusion criteria**
I1. Studies the main focus of which is on phishing (defined in Table 1.1) education, training and awareness. **I2.** Articles containing information about the research questions of this MLR.	**E1.** Studies that focus on automated phishing solutions (as defined in [77]) to counter phishing. **E2.**The study is a literature survey or review. **E3.** Full text of the study is not available. **E4.** The study is a short paper of fewer than 6 pages. **E5.** Studies not written in English. **E6.** Studies that have CORE rankings less than B. **E7.** Studies with CORE rankings of B and published before 2012.

We refined our search queries by conducting two preliminary searches. In the first search, we reviewed highly-ranked studies (CHI, SOUPS, IEEE S&P) and search queries used in previous surveys (discussed in Section 2.2) related to the scope of our research. To develop our search queries, we used the term "phish*" consistently and combined it with alternative terms from a predefined list that reflected our study's focus. These alternative terms included "train*", "awar*", "educat*", "teach*", "learn*", "interven*", "nudge*", and "warn*".

In the initial phase of our search, we found that the above-mentioned keywords did not generate sufficient user-based studies that offered guidance or assistance to users in combating phishing. To include potential relevant studies, we conducted a second pilot search. This phase of the search was more complex due to inconsistent terminology used in the literature, as explained by Franz, Zimmermann, Albrecht, *et al.* [21]. To identify studies closely aligned with the scope of our research, we used keywords such as "security indicators", "alert*", and "protect*", which resulted in more comprehensive and relevant findings. We examined the titles, abstracts, and keywords of previous studies to perform our search.

Grey literature

We choose Google[2] as a search engine to collect grey literature, like numerous prior studies [88]–[91]. Given the voluminous nature and complexity of grey literature in comparison with academic literature, we streamlined our search keywords when searching grey literature by including only the search terms "education", "training", "awareness" alongside the search term "phish*" when collecting grey literature, to produce more targeted and relevant results. We will explain why we use more general keywords for the grey study collection in the next paragraph.

[2] https://www.google.com/

Scopus permits searching documents based on title, abstract, and keywords. This allowed us to use multiple common terms and synonyms linked to education, training, and awareness within the search string to retrieve more targeted academic literature. However, Google's method of searching is different from scholarly databases like Scopus. Google searches the specified search terms across all indexed pages [92]; As a result, we observed that it yielded irreverent outcomes when the same set of keywords employed for Scopus was used for the Google search. For instance, when using the search terms "protect*" and "learn*", Google returned outcomes on ways end-users can protect themselves from phishing (e.g., [93]), which were not relevant to our study's objective of examining the challenges and critical success factors of diverse anti-phishing interventions. Hence, following the guidelines of [83], we decided to use a separate approach for coming up with the search string for Google. To come up with the search string for Google, we explored grey literature from our pilot study to identify relevant terms that discuss anti-phishing interventions. A similar approach has been taken in previous studies for grey literature search (e.g., [90]) due to the differences in searching in Google than other scholarly databases.

Our search keywords for both academic and grey literature are shown in Table 2.2.

2.3.3 Study selection

A search was performed on the 2nd of May 2022 on Scopus, which returned 2760 articles of academic literature. During the execution of these searches, we did not restrict the searches by using any filter (e.g., time limit, type of publication, publication venue) to ensure the comprehensive coverage and collection of the relevant articles. We then applied the inclusion and exclusion criteria, as detailed in Table 2.3, to remove the articles that are irrelevant to the scope of this study.

The studies selected for this study were chosen based on their primary focus on Phishing Education, Training and Awareness (PETA), as defined in Table 1.1. Specifically, we included studies that examine how to improve users' ability to combat phishing attacks or investigate strategies to help users detect phishing attempts (inclusion criterion I1). It should be noted that this differs from a body of research that investigates users' susceptibility to phishing attacks. In our pilot study, we encountered grey studies that discussed PETA but lacked relevant information regarding the challenges (RQ1) or critical success factors (RQ2) of PETA. Therefore, we established inclusion criterion I2. As one of the goals of this study is to contribute to enhancing the *usability* of PETA interventions, we excluded studies that discussed anti-phishing solutions that do not require user intervention or users cannot see or act upon (automated solution), as indicated by exclusion criterion E1.

To avoid including low-quality papers, we adopt a quality assessment approach based on publication venues. We identified the CORE ranking[3] of each publication venue of our search results and excluded papers with rankings below

[3] http://portal.core.edu.au/conf-ranks/, http://portal.core.edu.au/jnl-ranks/

CORE B. The CORE ranking is a process for ranking the academic Journals and Conferences in Computer Science (and related areas) that incorporates both expert domain knowledge and empirical data. The CORE ranking committee performs a thorough analysis of various factors, such as the citation count of papers published in the venue, the extent of involvement of leading researchers, as determined by metrics such as the author's h-index, the acceptance rates of the venues, as well as the expertise and engagement of the Program Chair (PC), as gauged by metrics such as the PC's h-index to rank the journals and conferences. The ranking system is regularly updated and refined to remain responsive to the changing needs and trends of the academic community [94], [95].

The CORE ranking is considered to be relevant outside of the Australian academic context because it utilizes widely recognized data and methodology, as mentioned above, for evaluating the quality of academic Journals and Conferences. Using empirical data and expert domain knowledge ensures that the ranking process is objective, transparent, and reliable, thus further enhancing the credibility and relevance of the ranking system beyond the Australian academic context. Furthermore, the international academic community has widely accepted and recognized the CORE ranking system, as CORE rankings being used as a benchmark for study selection in many existing systematic literature reviews published in leading Journals and conferences (e.g., [21], [96], [97]).

To achieve an integrated and comprehensive analysis of up-to-date research on PETA, we removed papers with CORE rank B published before 2012. The primary rationale was to eliminate outdated studies to report challenges (RQ1) and critical success factors (RQ2) pertinent to current intervention design practices. Nevertheless, despite being published before 2012, we retained five CORE A* and A ranked studies. This decision was justified by the fact that these studies provided recommendations that are relevant and applicable in the present context, for example, our included CORE ranked A* study published before 2012 [35] provides suggestions like "providing clear choices for better understanding" and "interrupting users' primary task to draw attention" for browser phishing warnings. These suggestions are relevant in the present context regarding the browser type and versions [35]. We have not considered h-index and citation count for inclusion/exclusion like some existing studies (e.g., Croft, Xie, and Babar [97]) as the calculation of both of these metrics relies on citation count. Consequently, it is difficult to assess the quality of recently published studies (e.g., those published within the last three years) based on these metrics [98].

Our Google search on the 22nd of April 2022 resulted in 121,000,000 items of grey literature. A large number of search results was attributed to the method of Google's search algorithm, which searches for specified terms across all indexed pages [92]. Moreover, these results contain academic studies and duplicate websites. Based on a careful examination of the search results, it was observed that results from page 17 to a few pages onward were either irrelevant or repetitive (we investigated up to page 20). As Garousi, Felderer, and Mäntylä [83] suggested to stop the search when no relevant or additional information is found, we decided to limit our analysis to grey literature from pages 1 to 16 of the

TABLE 2.4. Data quality assessment checklist for grey literature

Criteria	Questions	Score
Authority	**Q1.** Is the source from a reputable organization?	Yes - 1, partly - 0.5, no - 0
	Q2. Has the author published any other article in the area?	Yes - 1, partly - 0.5, no - 0
	Q3. Is the author an expert in the area?	Yes - 1, partly - 0.5, no - 0
Methodology	**Q4.** Does the source clearly state the aim?	Yes - 1, partly - 0.5, no - 0
	Q5. Is the source supported by credible references?	Yes - 1, partly - 0.5, no - 0
Date	**Q6.** Is the publication date clearly indicated?	Full date/year only/ month and year - 1, day and month only - 0.5, no date - 0
Novelty	**Q7.** Does the source support or oppose a current position?	Yes - 1, partly - 0.5, no - 0
Outlet type	**Q8.** What is the source's outlet type?	1st tier - 1, 2nd tier - 0.5, 3rd tier - 0

Google search results. Duplicate information from the same source shown on different pages was considered redundant and excluded. Academic studies from Google search results and duplicated Scopus results were also excluded. We further crawled through each link included in the websites of these 16 pages to collect further relevant studies. After reading the title, objectives, and the full article contained in these 16 pages, we retrieved 37 items of grey literature from Google.

2.3.4 Article quality assessment

When conducting data quality assessment, it is vital to follow the predefined criteria to ensure an unbiased collection process for primary studies [99]. Within the context of software engineering systematic literature reviews, quality assessment is conducted mainly through (1) explicitly defining assessment criteria and extracting them from primary studies, or (2) establishing research questions or inclusion/exclusion criteria that address quality concerns [100]. In our study, we followed a similar approach mentioned in the second category. We defined a CORE ranking-based study selection in the inclusion/exclusion criteria and opted to conduct a quality assessment for academic studies. We contend that the CORE ranking-based study selection method facilitated the selection of high-quality papers for our study. This is primarily due to two reasons. Firstly, as mentioned before, CORE ranking is a rigorous process that involves experienced committee members ranking conferences and journals based on several widely accepted evaluation metrics (e.g., citation counts of papers published in the venue, author's h-index, acceptance rates of the venues) [94], [95]. Secondly, the Reviewers and Program Committees (PCs) of these venues are composed of experts from both academia and industry who perform peer reviews based on several metrics (e.g., novelty, correctness, contributions with well-supported

methodology, impact, reusability, practicability) to ensure the scientific validity of the results [101]. The significance of this aspect is particularly pertinent in the context of our study. Our research aims to contribute to the enhancement of current and future anti-phishing interventions by identifying challenges (RQ1) and critical success factors (RQ2) through the synthesis of relevant literature. For this reason, the selected studies should adhere to a rigorous and established methodology to ensure that their findings are credible and robust. Furthermore, by adopting a CORE rank-based study selection method, we were able to circumvent the subjectivity involved in defining and scoring the quality assessment, as noted in a previous study [100].

In contrast, as grey literature is non-peer-reviewed, we adopted a rigorous quality assessment checklist suggested by Garousi, Felderer, and Mäntylä [83]. Table 2.4 shows the criteria and the corresponding scores for each criterion. We have selected eight assessment criteria related to the reputation of the published authority (Q1), author of the publication (Q2), authors' expertise (Q3) etc. The responses related to the eight criteria were classified into three categories: "Yes", "Partly", and "No", with corresponding scores of 1, 0.5, and 0, respectively, as adapted from the previous studies [102], [103]. In contrast to the scoring process recommended by Garousi, Felderer, and Mäntylä [83], where a score of 1 is assigned to "Yes" and 0 to "No", we deviated from this approach due to the variability in our data. Specifically, for criteria Q6 and Q8, we encountered three different types of data (discussed in the following paragraphs) that necessitated distinct scores to facilitate an accurate cumulative assessment similar to the other criteria.

To test the validity of our scoring system and to fine-tune the scoring metrics, a pilot study was conducted on a randomly selected sample of 16 articles from the grey literature. During the application of Q6, discrepancies were observed in the availability of publication dates for the articles. Some articles lacked any date information, while others only provided the year or specific day and month details. Given that the year signifies the recency of an article, it was considered the most crucial component of the date. Consequently, a score of 1 was assigned to articles with a known year of publication, regardless of the presence of day or month information. On the other hand, a score of 0 was assigned when no date was available. Articles that only provided the day and month received a score of 0.5 in Q6. Regarding Q8, aside from the outlet types (tier 1, tier 2, and tier 3) defined by Adams, Smart, and Huff [104] and later refined by Garousi, Felderer, and Mäntylä [83] in the guideline, three additional article types (research reports, case studies, and guides) were discovered during the search process. To address this, we employed our understanding and observations to map these articles to the corresponding outlet types as follows: research reports were categorized as tier 1, case studies were categorized as tier 2, and guides were categorized as tier 3.

From the pilot study, we also observed that every grey article in our pilot sample scored a minimum of 5. The substantial variations in scores are mainly attributable to variability in criteria Q5, Q6, and Q8. As most of the articles in our pilot sample scored at least 0.5 for Q5, Q6, or Q8, we have chosen a cut-off value of 5.5 to exclude low-quality articles at this quality assessment

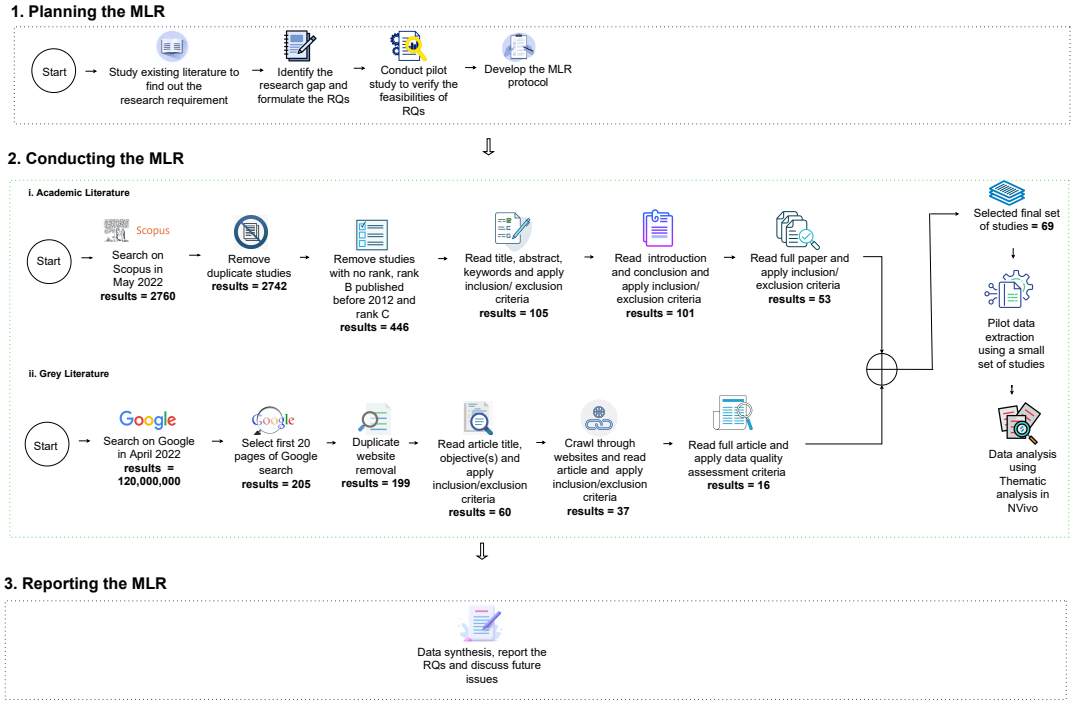

FIGURE 2.1. Pictorial representation of our research methodology

stage (demonstrated in Equation 2.1).

$$\text{Quality assessment score}, \sum_{i=1}^{8} Q[i] > 5.5 \tag{2.1}$$

Here i is the question number mentioned in Table 2.4 and the value $Q[i]$ is either score 1, 0.5 or 0. Applying this process consistently across all items of grey literature resulted in the selection of 16 articles.
Our list of primary studies can be found in Appendix A and Appendix B.

2.3.5 Data analysis

An overview of our data extraction and data synthesis process is provided in this section.

Data extraction

We systematically prepared and refined a data extraction form (please refer to Appendix C to see the data extraction form) to collect different types of data used for this study by following the existing guidelines [99], [105]. Apart from collecting data for reporting our formulated research questions, we collected demographic information to gain a general understanding of the data included in this study (e.g., distribution and trends in the number of articles over time). We also collected limitations, threats to validity, and recommendations for future research directions. As suggested in the guideline [99], data extraction was performed by more than one author (in our case, two). The first author extracted data from 49 studies and 14 grey studies, while the second author

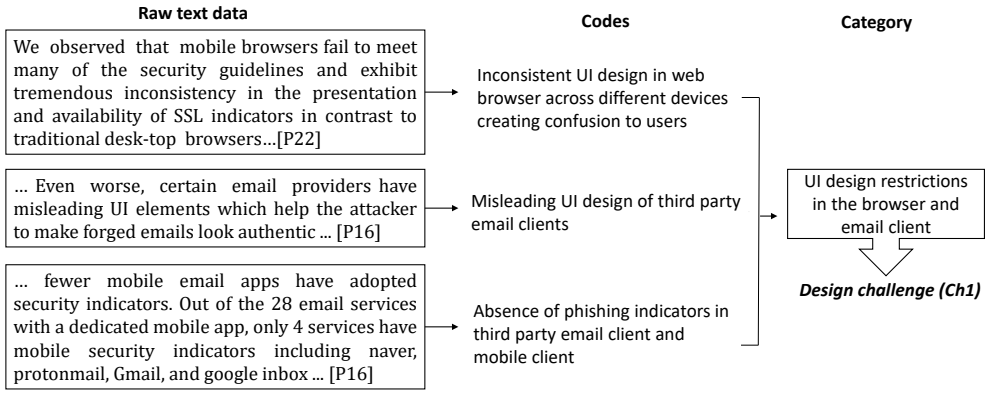

FIGURE 2.2. An example of our data analysis process

collected data from 4 studies and 2 grey studies. The data extraction form in Excel format was uploaded to the shared folder and discussed in the weekly research meetings among all the authors.

Data synthesis

The raw text collected from the primary studies, in accordance with the research questions guiding the investigation, was unstructured, encompassing a diverse range of information that can be challenging to interpret effectively. As a result, to gain a more comprehensive understanding of the significance and contextual relevance of the data, as well as to identify recurrent patterns that can facilitate the resolution of research questions, we conducted a thematic analysis. This method enables the exploration of the intricacies and subtleties of unstructured data by analyzing the raw text data in a systematic and rigorous manner [102]. Consequently, the use of thematic analysis allows for the elucidation of insights into the complexities of unstructured data.

We adopted the thematic analysis process discussed by Braun and Clarke [106]. Our thematic analysis process was conducted on Nvivo, a tool for qualitative data analysis [106]. Our extracted data, stored in an Excel spreadsheet, was imported into Nvivo, and then *open coding* was performed using this tool. *Open coding* is a process where labeled data (usually referred to as *code*) is obtained by breaking down the data into small components [107] and labeling each component. We performed the open coding process through continuous iterations of extracted data (i.e., codes generated in the initial stage were modified and updated in later stages).

Initially, a pilot data extraction was performed by the first author with a set of five academic studies and two grey studies (randomly selected) to understand the pattern of the data. We scrutinized all the codes and grouped them into *themes* based on the similarities of the codes by utilizing the multi-layer structure of Nvivo. The first author revised the themes after weekly research discussions, and any suggestions or feedback from other authors were incorporated accordingly into the data analysis process. Figure 2.1 demonstrates an overview of our research methodology and Figure 2.2 displays an example of our data analysis process. The main findings are discussed in Section 2.5 and 2.6.

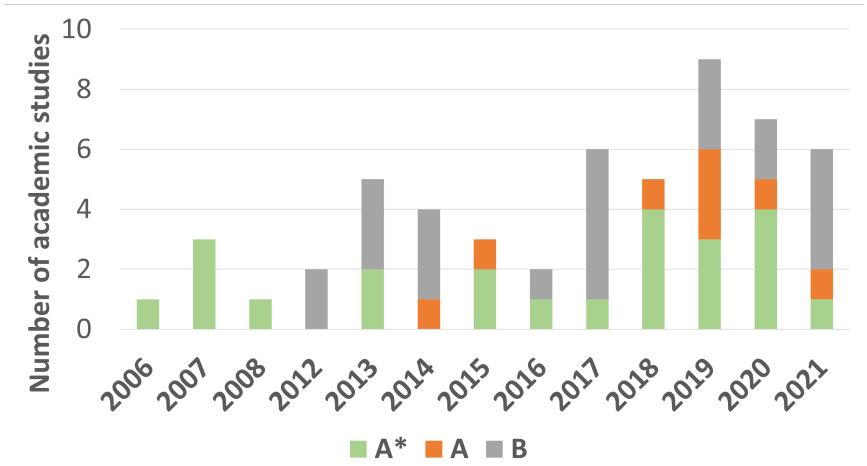

FIGURE 2.3. Number of academic studies over years and CORE ranking

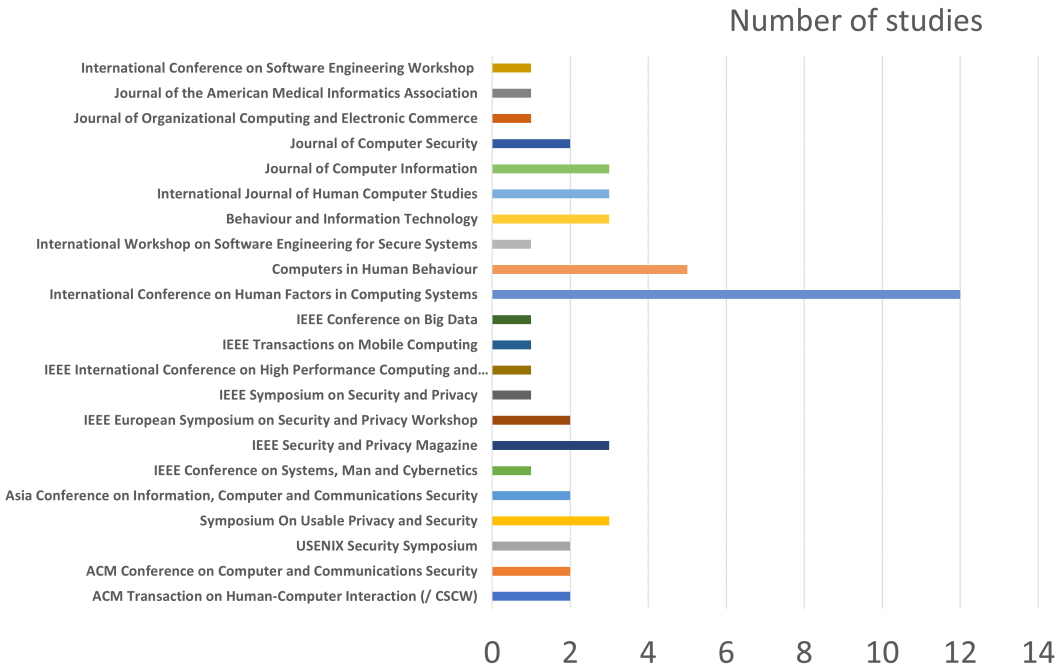

FIGURE 2.4. Distribution of academic studies over type of venues

2.4 Demographic Description of Selected Studies

We report demographic data in this section in relation to PETA, such as the main publication venues for this area of research and the most/least investigated interventions. An overview of demographic data can help new researchers gather useful information about the domain [84]. Some key insights about the demographic data are provided below:

• PETA first began drawing significant research attention in 2006. This domain has experienced rapidly increasing popularity in the last 4 years (2017-2021) (Figure 2.3). Starting from 2006, our pool of primary studies includes no

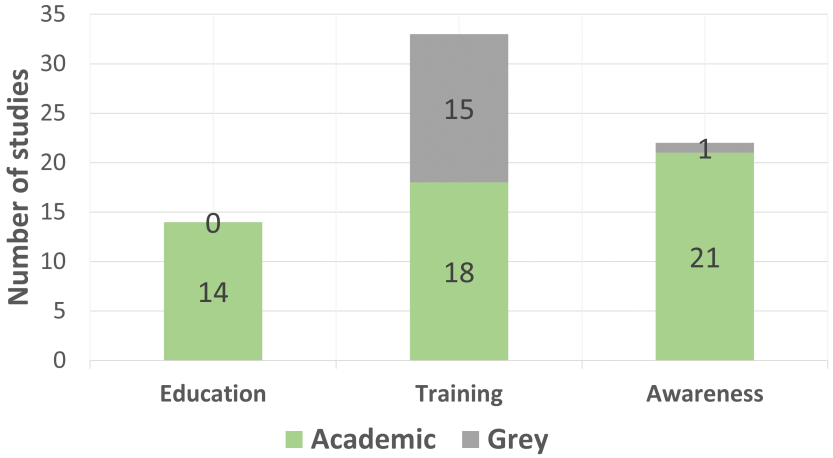

FIGURE 2.5. Number of primary studies for each phishing intervention

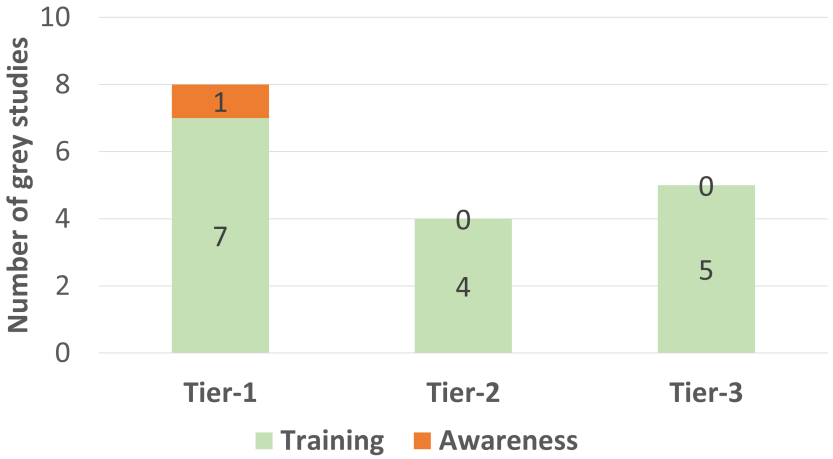

FIGURE 2.6. Number of grey studies over tier types and phishing intervention

academic study from 2009 to 2011. Our pool of studies contains only A*-ranked publications from the year 2006 through 2008. Also, as we have decided to omit CORE B-ranked studies published before 2012 during our inclusion/exclusion stage, we do not have any B-ranked studies included before 2012. At the time of our search, we have not found any high-ranked (Rank A*, A or B) study published in 2022.

• Figure 2.4 represents the venues of our selected academic studies ranging from journals, conferences, magazines, and workshop papers. International Conference on Human Factors in Computing Systems (CHI) is the most popular venue for publishing research in this field. Our pool of studies contains 22% studies (12 academic studies out of 53 academic studies) from this conference. Human-centric and security-related publication venues appear frequently; nevertheless, research outlets with other focuses, such as biomedical engineering and health sciences (e.g., Journal of the American Medical Informatics Association), also appear frequently, indicating that phishing is a common concern

shared by researchers and practitioners from across different disciplinary fields.

• Figure 2.5 demonstrates that we have more studies (both academic and grey) related to training, compared to education and awareness interventions. This data also indicates that education intervention is the least explored category out of the three. The total number in Figure 2.5 does not add up to 16 (number of grey studies) as one of the grey studies (P54) falls under both the phishing training and phishing awareness intervention categories. This dual categorization is attributable to the fact that the mentioned grey study contains information pertaining to both anti-phishing tools (which aligns with the phishing awareness category) and phishing training (which is associated with the training category).

• The type of grey studies we collected include whitepapers (tier-1, 43.75%), annual reports (tier-2, 25%) and blogs (tier-3, 31.25%) displayed in Figure 2.6.

2.5 Challenges in the Design, Implementation and Evaluation of Anti-phishing Interventions

This section provides a holistic overview of the limitations and obstacles facing in the design, implementation, and evaluation stages of anti-phishing interventions.

2.5.1 Challenges in the design

Design limitations broadly cover limitations associated with the design and performance of anti-phishing interventions. The reported challenges mainly focused on the currently missing features in the design of phishing interventions.

Challenge 1. UI design restrictions in browsers and email clients

Design consistency in user interfaces is one of the critical usability attributes. A consistent design across different interfaces provides users with many benefits, including enabling them to transfer knowledge and skills across other similar systems, hence reducing their time and effort spent in learning to use the new systems [P22, P49]. A study documented that browser designers follow the same web-security guidelines for designing user interfaces for both mobile and desktop browsers [P16]. This leads to *inconsistent UI design*, which creates confusion among users and increases their risk exposure to phishing. The main reasons behind inconsistent designs are the lack of communication between mobile and desktop developers and the choice that developers need to make between usability and security. Mobile browsers have small display sizes; additionally, the padlock icon and the HTTPS URL prefix indicators in the address bar are hidden to accommodate the contents on a small display. Therefore, to make the content visible to the users and to keep a clean interface, less important information is often removed by the developers. Unfortunately, phishing indicators

are one of the pieces of information removed from mobile display in this process. Sometimes it is cumbersome for users to view the address bar in order to inspect the phishing indicators, which leaves phishing warnings unnoticed and exposes users to greater risk of phishing attacks.

Additionally, most mobile browsers do not use phishing indicators (e.g., extended validation SSL indicators). Many email providers and third-party email clients (Microsoft Outlook) do not provide warnings for forged emails. Moreover, some email providers (e.g., Gmail, Yahoo, Apple iCloud) incorporate *misleading UI design*, which makes the spoofed email appear legitimate, for example, by employing confusing colors to indicate the legitimacy of an email [P16]. This issue arises due to miscommunication between email providers and end-users. Furthermore, the *absence of phishing indicators* grants potential attackers the opportunity to pass off the phishing email as a legitimate email.

Challenge 2. Content restrictions for phishing education and training

The effectiveness of phishing training largely depends on the content of the training material [P7]. Existing phishing training is *less engaging* as it allows one-way transmission of knowledge and does not provide immediate feedback to the users [P10, P19, P28]. Moreover, training materials often fail to capture the interest of users (e.g., a serious gamer[4] will not be interested in playing educational games designed for casual gamers) [P36]. Some educational games have *complex interface and configurable scenarios* in the content, which renders them unsuitable for shorter training periods (e.g., CyberCIEGE) [P28]. Attackers use diverse attack vectors and a wide range of deceptions to lure users into disclosing their information. Unfortunately, existing training materials cover only *limited attack vectors* in the content. Consequently, users may only learn about some aspects of phishing attacks but not others, for instance, by learning about how to detect malicious URLs or deception cues, but remaining unaware of malware attachments that may also be enclosed in the email [P19, P24, P59]. Often training emails are *lengthy and wordy*, making it more time-consuming for users to make decisions. As a result, some users become confused about what messages the training email is trying to convey. This *time-consuming decision-making process* causes several problems to the users, including distraction and fearfulness, given that each user receives multiple emails per day [P5, P7]. Current designs of anti-phishing education materials do not consider users' *knowledge gap and misconceptions*, for instance, little consideration is given in the design of phishing training to human-centric factors such as users' state of mind (e.g., factors driving users to insecure behaviors and poor decisions making), how scammers operate, how users can be targeted, and users' strategies for dealing with phishing risks [P11, P19]. Contents containing *repetitive information* [P7] which users already have seen previously and the *presence of cultural bias* [P36] in the content (e.g., content on specific language, URLs from websites of a specific country) are also significant limitations in the design of current training contents. Repetitive training content does not add value to the knowledge and consumes users' valuable time. Biased training content is only

[4] "Serious games are a set of solutions developed to make sessions more fun and less boring. A generally accepted definition is - video games developed for a primary purpose other than pure entertainment" [108]

effective when users are familiar with the given information (e.g., websites or URLs from a particular country).

Challenge 3. Design constraints for anti-phishing warning UI interfaces

Several phishing indicators have been developed to warn users about phishing, including browser security toolbars such as SpoofGuard [109], Trustbar [110], and site authentication image (a user-defined image selected for login, which enables users to verify the image before entering personal information) [P3]. Unfortunately, due to the *design similarity of phishing warnings with other security warnings* [P1, P28] and *lack of active interruption* [P1, P11, P14, P43, P44], phishing indicators often fail to attract user attention and lead to habituation to such warnings. A more salient function is needed to interrupt user actions in order to prevent users from succumbing to phishing attacks. Moreover, *frequent exposure to the warning* decreases users' neural activity, leading to carelessness, laziness, warning fatigue, habituation and disturbance. As a result, users tend to automatically ignore phishing warnings [P4, P13, P14, P17, P18, P26]. *Unsuitable warning placement* is another reason why warnings go unnoticed. Most phishing indicators are small, making it difficult to attract user attention. For instance, in a phishing email where a malicious URL is the main hazard, current warnings are placed too far from the malicious links, which provides limited help to the users in identifying phishing URLs. The warnings are not visually prominent and are not easily noticeable by the users [P2, P3, P5, P7, P11, P15, P25].

Challenge 4. Problems with anti-phishing warning content

Some email providers use warning banners to indicate phishing risks to the users. Such banner warnings typically do not provide sufficient explanation or reasoning to enable users to identify or assess phishing risk (for example, typically, no explanation about why a link is deemed malicious is provided in the banner warning). The displayed information needs to be more comprehensive to enable users to make informed decisions and to enhance trust in the warnings. Although some current warnings offer a "learn more" button, they typically only contain general advice. The absence of specific information places an extra cognitive burden on users to locate the suspicious cues in the links. Instead of providing concrete information, current warnings are made unnecessarily *lengthy*, which is time-consuming to read [P41]. A *lack of comprehension* of the warning content, exacerbated by the absence of any justifications, leads to misunderstanding or ignorance [P14, P18, P25, P49]. Most often, the warning content is designed to target security-conscious and experienced users, ignoring the needs of novice or non-expert users. Additionally, the *lack of consistency in design practices of security warnings*, which vary significantly across different vendors, platforms, and browser versions, further creates confusion among users [P49].

Challenge 5. Performance limitations of anti-phishing tools

Anti-phishing indicators are designed to provide accurate information to users by accurately identifying potential sources of phishing. However, studies have shown that existing tools suffer from limitations in their usability and performance, which potentially impede the effectiveness of current anti-phishing tools in helping users detect phishing. The main source of usability issues arises from the tools providing inaccurate information to users by incorrectly identifying or failing to identify suspected phishing attempts (i.e. generating false positive or false negative results), which can further result in users' distrust of these tools. The current literature offers limited evidence to show that anti-phishing tools are indeed protecting users from phishing attacks. It remains unclear how or to what extent these tools assist users in determining a website's legitimacy. Designing usable phishing indicators remains an unresolved problem in the usable security domain [P1, P2, P8]. Specifically, *False positive results* are an important limitation of existing anti-phishing tools. Some common reasons for falsely identifying legitimate sources of phishing attempts include: 1) delay in updating lists of known phishing sites, 2) improper maintenance of whitelist, 3) inaccuracy in performing successful detection (even the best phishing indicators miss 20 percent of phishing websites), 4) spyware infection in user device causing failures of site authentication image (e.g., Site-Key), which is used to protect user password during login to the page but cannot function in protecting users' personal information if users' computers are infected with spyware. Consequently, users learn to distrust and ignore the phishing warnings if previous warnings have mistakenly provided incorrect information (e.g., showing that a website is phishing when it is legitimate, displaying a phishing indicator even when there is no phishing risk, absence of phishing indicator when there is a high risk of phishing) [P1, P2, P3, P8, P10, P13, P14, P18, P24, P25, P28, P44, P49, P57, P69].

Challenge 6. Lack of attention to phishing indicators

For phishing indicators to serve their purpose, they must be heeded by end users. However, evidence shows an alarming frequency with which end-users ignore phishing warnings due to several human-centric causes. First, users are more likely to ignore security warnings about possible phishing attempts during their online activities if they do not understand the risks and consequences of phishing attacks. This lack of knowledge results in a reduced likelihood for users to pay attention to phishing warnings [P1, P2, P3, P4, P8, P11, P14, P24, P28, P31, P36, P39, P44, P49]. This problem is compounded by the fact that many users misunderstand the nature of anti-phishing toolbars, as many users mistake the browser toolbar for an advertisement banner and are unsure if the toolbar is brought up by the browser or the website they are visiting. This lack of knowledge further makes it more difficult for users to interpret what a phishing warning is trying to convey. For example, prior evidence shows a lack of understanding by users of the nature and information conveyed by the anti-phishing toolbars such as the Neural-Information toolbar [P2]. Furthermore, a lack of knowledge about phishing can lead some users to apply the wrong

anti-phishing strategy and distrust the security warnings (for instance, if users believe their approach to identifying phishing is correct, they would distrust any incongruous phishing warnings as wrong) [P2]. Another common reason why users fail to heed phishing warnings is when users have misplaced *"confidence in the websites they visit"* [P2, P8, P14, P24, P49]. For instance, users may cognitively deem a website to be legitimate due to its *look and feel*, causing users to ignore phishing warnings that show "proof of authenticity of a website". Users also sometimes exclusively rely on one specific phishing indicator (e.g., site authentication image or SSL indicators) to the exclusion of others, causing users to ignore important information offered by other indicators, which can contribute to determining the authenticity of a website [P3, P8]. Familiarity with a site or brand can induce users to trust the website [P11], thereby overlooking phishing warnings. For example, users tend to trust the website that they had previously visited. Moreover, reusing personalized phishing indicators allows attackers to develop an attack vector in the hope that users will use the same indicators for different applications [P31].

Challenge 7. Need to design specific training for spear phishing

Spear phishing is a type of phishing attack where attackers use victims' personal information to initiate the attack. Training users to detect spear phishing attacks is crucial as they are more effective and cause greater harm than regular phishing attacks, due to the level of personal relevance involved in phishing emails [P49]. However, the literature has identified numerous challenges that render it difficult to train users for spear phishing [P26]. Examples include emails mimicking standard business processes [P58], emails from known organizations [P1], the timing of email receipt [P1], previous simulated phishing emails sent internally by organizations [P7], emails that do not request any personal information [P7], messages from trusted sources (e.g., friends) [P15, P21, P44], and click-whirr response tendency (automatically responding to repeating events) [P44]. Due to the aforementioned reasons, users often ignore phishing warnings [P1] and face difficulty recognizing spear phishing emails [P7]. Often personal relevance invokes curiosity among users to click on a malicious link [P15]. The factors exacerbate users' vulnerability to spear phishing attempts, which will continue to be successful unless more attention is paid to the design of interventions aimed at training users to detect spear phishing attacks.

Challenge 8. Disregard for users' mental limitations during design

Human behaviors and decision-making are non-deterministic and unpredictable. A user who is security-conscious one day may act differently the next day due to some human-centric factors, such as illness or attention overload, resulting in greater exposure to security risk [P24]. Human-centric vulnerabilities constitute a greater risk factor than technical vulnerabilities to allow attackers to breach system security more easily [P24], as no complex cryptographic knowledge is required on the attackers' part to exploit human-centric vulnerabilities in phishing attacks [P49]. Despite the central role played by human perception and information processing in user decision-making, there remains a surprising

lack of effort to incorporate users' cognitive constraints into the design of anti-phishing tools [P24]. Incorporating users' cognitive constraints into the design can be challenging. Unlike machines, human vulnerabilities cannot be patched with straightforward solutions, as human behaviors cannot be regulated, controlled, or changed by technical modifications [P15, P40]. Current anti-phishing technology fails to take into account some important human behavioral factors in the design process. For instance, 1) security is not the primary concern of users, whose attention is usually focused on other online tasks they are performing (such as reading books, checking email, or making online purchases), 2) users do not check security notification continuously [P1], 3) users cannot attend to everything at the same time due to cognitive limitations [P13], 4) when occupied with tasks that dominate their attention, individuals often fail to notice *"highly conspicuous but unexpected"* events [P13, P47], 5) reliance on *warning disruptions* can cause users to undertake a more passive view towards security consciousness and take no active actions to avoid phishing attacks [P13]. In summary, besides displaying phishing warnings to the users, there is a significant need to provide users with an *alternative options to complete their task* (e.g., suggestion for alternative website). Otherwise, warning disruption may cause some users to take risky actions to achieve their goals. Overall, the current lack of consideration for human-centric factors, such as users' cognitive constraints and lack of motivation or attention, in the design of anti-phishing tools and warnings can significantly restrict the effectiveness of these tools [P24].

2.5.2 Challenges in the implementation

Implementation challenges represent obstacles in adopting, deploying, and automating anti-phishing technologies, as well as weaknesses of the current policies and guidelines regarding anti-phishing defense.

Challenge 9. Anti-phishing technology deployment challenge

Existing studies have proposed small prototypes of anti-phishing tools or techniques (e.g., browser plug-ins) for the purpose of easy deployment on different platforms in real-world settings [P23]. However, real-world deployment of anti-phishing technologies faces several challenges. First, organizations outsource phishing awareness and training material, however, it is often managed by internal staff members, which poses challenges to managing the materials optimally (optimal management is a shared responsibility). The involvement of external service providers to provide support on phishing training program content or tool development makes it difficult to keep track of the changes and to measure the effects of phishing detection capability. Second, it is difficult to make employees understand, identify, and safeguard their personal information due to a common misconception that ensuring security is only an IT problem rather than a responsibility shared by all personnel of an organization. Third, distributed/siloed work environments and expanded infrastructure (e.g., new vendors, SaaS applications) create an enlarged attack surface, rendering it difficult for IT personnel to navigate as such navigation requires a team effort and coordination among team members. Fourth, employees who work from home

might lack sufficient infrastructure support. Moreover, the absence of adequate technical skills to set up a secure personal computing system increases their phishing risks [P6, P54, P57, P62, P65]. Fifth, it is difficult to configure email clients to allow phishing training emails to be delivered to mailboxes, as training emails are often classified as spam by automated email security countermeasures [P28]. Finally, deploying anti-phishing tools and anti-phishing browser plug-ins can be a complicated process due to the interdependency of technological, organizational, individual, and procedural factors, such as browser platform dependency (e.g., Mozilla Firefox, and Internet Explorer) and choices of IT standards or frameworks (e.g., Information Technology Infrastructure Library (ITIL)/Control of Business Objectives and Technology (COBIT)). Security service providers usually provide only high-level guidance, while ignoring the need for minor decision variations in specific organizations [P23, P38, P50].

Challenge 10. Technology adoption and usage challenges

Anti-phishing interventions (e.g., personalized security indicators, educational games, third-party anti-phishing tools, and anti-phishing training interventions) are developed to help users identify and assess phishing risks. Yet several usage-related challenges continue to impede the widespread usage of these anti-phishing interventions and hamper their success. These challenges include: 1) some tools are difficult to use by non-expert users, 2) personal security indicators require extra efforts during installation and when using the applications [P37, P45], 3) the success of anti-phishing training depends on users' willingness and ability to learn and recall the information, and their capability to apply the learned information in subsequent situations [P19, P31, P45], 4) adoption of game-based education requires a level of prior knowledge and investment of time and efforts to build the software [P37, P45].

Challenge 11. Challenges due to complicated URL and domain name structures

Although users can be trained to improve their ability to identify malicious URLs, doing so is increasingly difficult due to the complex visual traps and textual manipulations employed by phishing attackers (e.g., hidden links, crafted texts, additional texts). Studies have shown that, even after training, it was difficult for users to detect small URL changes (e.g., swapping of two letters) [P45, P46, P47]. Complicated URL and domain name structures give rise to user confusion. Identifying well-concealed cues from visually inspecting URLs demands a great deal of attention and effort from users. Minor discrepancies can go unnoticed if user attention is distracted [P46]. This problem is exacerbated by the fact that some organizations lend their names for use by external parties: for instance, trustworthy websites, such as Microsoft, may provide hosting services for external web content (e.g., thereby enabling someone to pay Microsoft for web space and to create a website named *malicious.windows.net*). The resultant page is linked to a real Microsoft domain but its contents are controlled by the attackers.

Challenge 12. Obstacles to automating phishing incident response and anti-phishing training

Phishing attacks can cause damage within a few seconds [P50]. Therefore, it is essential to take prompt initiatives in response to phishing attacks. Organizations often run phishing simulations to train their employees or to test the ability of the employee to detect a phishing attack. Large phishing simulation campaigns can overload the "help desk" with phishing reports, which can impact on the regular workflow of staff and hinder the effectiveness of phishing mitigation processes. Automating the ticketing system and phishing incident responses can help enable organizations to respond to phishing incidents promptly and save administrators time and effort. However, the task of automating incident responses is challenging because the performance of the incident response is mainly determined by the accuracy of the initial report, which in turn requires manual confirmation and validation by experts [P50]. Additionally, automating phishing training is also challenging as the content of a training email needs to be manually written and crafted by administrators to make it more realistic [P45].

Challenge 13. Exploitation of software vulnerabilities by attackers

Anti-phishing plug-ins (e.g., Anti-Phish, Spoofguard) inhibit the transmission of user-sensitive information to the attackers' site by checking user inputs containing sensitive information. Nevertheless, such anti-phishing plug-ins are not foolproof as they are less effective when attackers use malicious JavaScript on their phishing websites. Such JavaScript provides attackers with opportunities to bypass monitoring phishing plug-ins. For instance, the use of JavaScript allows an attacker to listen to a critical press event on the client side and send each character back to the attacker's server before the user can press the submit button. Hence, users' sensitive information can be transferred to the attackers before the plug-in can detect that sensitive information. A solution to mitigate JavaScript attacks is deactivating JavaScript on web pages that include forms. However, it is not feasible to do so across the board, as many legitimate websites use JavaScript for form submission [P23]. Similar to a Javascript attack, attackers can also exploit other software-based vulnerabilities, such as by employing cross-site scripting (XSS) to inject malicious codes into the login pages executing on the client side to steal users' personal data. In this way, even an experienced user can be deceived into giving away personal information as the webpage containing malicious codes refers to a legitimate webpage [P49].

Challenge 14. Unguarded email clients and websites

Simple Mail Transfer Protocol (SMTP) does not use any built-in mechanism to prevent email spoofing. It relies on the SMTP extensions - Sender Policy Framework (SPF), Domain Key Identified Mail (DKIM). and Domain-based Message Authentication, Reporting and Conformance (DMARC) deployed voluntarily by the email providers for authentication. SPF helps senders maintain a DNS record containing a list of authorized IP addresses, which are allowed to

send email from a specific domain. DKIM is a cryptographic signature used to sign an email to ensure that the email originates from an authorized source from that domain. In the absence of an SMTP authentication mechanism, attackers can send emails from spoofed email addresses. As the deployment of SPF and DKIM is not compulsory, few email providers have adopted them. It is unclear how email providers handle an email that fails the authentication process [P16]. Furthermore, many e-commerce companies do not use SSL to protect their login page. SSL certificates ensure that users' data is protected and verify the authenticity of a website. Poor security practices by email clients and websites make it more difficult for users to identify the legitimacy of an email or website [P2].

Challenge 15. Limitations of current anti-phishing planning, policies, and guidelines

Studies have uncovered significant flaws in existing anti-phishing guidelines, policies, and phishing training practices [P15, P42, P50]. Many organizations offer guidelines on their websites to enable their employees to learn how to identify phishing attacks. However, the information contained in such guidelines often contains generic information and does not include detailed explanations of the consequences of these attacks, nor information about potential advanced phishing techniques, such as clone phishing[5]. Moreover, contradicting information included in the guidelines creates confusion and disturbance among users. Contradictory information also reduces readers' self-efficacy and ability to detect phishing [P42]. Readers of this information often misinterpret the information in the guidelines as a comprehensive list of possibilities, rather than mere examples. Consequently, users would recognize only the phishing cues mentioned in the guideline. Some information is incomplete or outdated, which can cause *"security fatigue"*. The absence of succinct and correct information also increases the likelihood for users to miss vital information or place trust in unreliable information [P15]. Organizations need to do more than merely adopt email security policies and guidelines to precipitate behavioral changes [P50]. More formal approaches are also needed to learn from past experiences involving previous phishing incidents. Poor planning in phishing training can diminish their impacts [P50]. For instance, if employees are summoned to the break room for phishing training on very short notice, this will likely reduce the effectiveness of the training session. Also, issues in the workflow process are not well understood to find the right tool or software, and often customized or outdated tools are selected without properly considering the best fit for phishing incident response [P50].

2.5.3 Challenges in the evaluation

The effectiveness of anti-phishing initiatives must be regularly evaluated to identify weaknesses and facilitate continuous improvement. This section discusses

[5]A type of phishing where a previously received genuine email is cloned to a malicious email [111]

challenges encountered in evaluating the effectiveness and usability testing of PETA interventions.

Challenge 16. Lack of industry relevance in evaluation practices and settings

Many existing studies have methodological limitations which reduce the relevance of their findings to the industry. For instance, limited consideration of users of certain demographics (e.g., children) and the use of artificial settings (e.g., controlled environments) to observe user behavior induces sample bias, which in turn results in a lack of generalizability, as the findings and applications can not be extended to other contexts [P1, P7, P13, P14, P18, P21, P26, P30, P32, P35]. For example, the findings of a study conducted with participants using Google Chrome might not be comparable with those using Mozilla Firefox, and the findings of a study experimenting with university students or adults cannot be extrapolated to children or adolescents. Despite children's extremely high vulnerability to phishing risks due to their credulity and lack of experience, children are the single most overlooked demographic group in the literature [P21, P35], as most studies do not commonly take children into account. Similarly, metrics used for evaluation also lead to erroneous calculations, for example, drawing the conclusion of phishing simulation based only on the click-through rate of simulation emails [P32].

Challenge 17. Complications regarding data collection and replicating user experience

A study setting that does not adequately replicate users' real-life phishing experiences would result in findings with low ecological validity. Existing literature has reported several challenges regarding data collection and replicating users' real-life behavior, for example, 1) collecting data from children is challenging due to the difficulty of obtaining and maintaining their attention during the study [P21], 2) participants may be disinclined to disclose truthful information about their past incidents of falling victim to phishing attacks, out of embarrassment or impression management [P40], 3) it is difficult to replicate users' real-life behavior during a phishing attack for various reasons: experimental settings lack the necessary element of risk [P3]; studies are often conducted with the help of a role-playing scenario [P3]; and Not evaluating user behavior in their regular working environment can affect their responses [P3, P43], 4) reimbursement or permission to opt-in before conducting the study poses some challenges. For example, participants might behave differently due to these environmental factors that are only present in the study setting [P14, P21, P40, P48].

Challenge 18. Insufficient usability and effectiveness evaluation of phishing interventions

Due to the continuous changes in the design of phishing interventions, it is essential to evaluate the usability of these interventions to understand user requirements and incorporate them into the design [P4, P8, P13, P17, P18, P30, P37, P40, P41]. A usability study could contribute to the performance enhancement of anti-phishing interventions and improve user learning experience [P36]. However, the rapid speed with which updates are made to software components can render past evaluation information obsolete. For example, a study evaluating the usability and effectiveness of browser phishing warnings or anti-phishing toolbar would no longer be relevant when a new version of the browser is available, given the changes made in the new version [P4, P8]. Effectiveness evaluation of a method or technique utilized to teach users is also important to improve their learning experience. Sometimes, organizations use techniques (e.g., distributing leaflets) to inform employees about phishing without properly evaluating how effective this approach is in practice to teach individuals about phishing. Such effectiveness could be low due to human-centric factors: for example, people may not pay attention to phishing instructions if they are deemed irrelevant to them (e.g., people with no prior experience with phishing might not realize that such instructions are pertinent to them) [P13]. Moreover, existing studies reported that rigorous empirical investigations using different methods and variables are required, such as the impact of individual and organizational factors on training effectiveness [P40], the role of different phishing cues in the decision-making process of phishing detection [P41], the impacts of different email types or contents during phishing attacks [P30], expert and non-experts' decision-making processes about phishing attacks [P41].

Challenge 19. Lack of sophisticated quantification of phishing training outcome

Organizations often run phishing training simulations to test their employee's ability to detect phishing attacks. This is done by sending employees fake phishing emails, and employees' ability is subsequently measured using different performance metrics. The most common performance metric used for evaluation is the number of times users end up clicking on phishing links. By counting the number of clicks, informs organizations about which employees need access to anti-phishing instructions and training. However, security tools and third-party software have bots that can click on all the links in an email in a sandbox environment. This process is executed to ensure no malicious URL is in the email. The URLs clicked by *bots* can be misinterpreted as originating from a human user, which creates false positive results in the phishing simulation reporting. It can provide false insights about the organizations' security [P55].

The measurement of phishing simulation outcomes can also be influenced by users' offline conduct such as prairie dogging, which refers to the phenomenon where an employee receives a phishing simulation email and lets other employees know about it. This practice can also affect click-on rates and, consequently, the measurement of outcomes from phishing simulations [P15, P59].

Challenge 20. Lack of post-training user knowledge retention practice

Although phishing training can improve users' knowledge regarding phishing attacks, such impact is subject to decay over time [P13, P21, P40, P45]. Studies have shown that phishing awareness returns to the pre-intervention level within a very short period [P21]. Even if users develop the ability to detect phishing through phishing training, they will soon struggle to remember the relevant information if they do not apply the learned information in practical situations [P45]. Currently, the effect and duration of this knowledge are seldom investigated, resulting in a limited understanding of how long the impact of PETA initiatives can persist and how often users require re-training [P7, P31, P34, P54].

Table 2.5 presents the congregated challenges along with the main key points. In order to differentiate the data derived from grey studies, the P numbers associated with such studies have been distinctly marked in green.

TABLE 2.5. Challenges in anti-phishing interventions

Challenges	Key points (included papers)	#
Design		
Challenge 1. UI design restrictions in the browser and email client A	① Inconsistent UI design in web browser across different devices creating confusion to users [P22, P49] ② Misleading UI design of third party email clients [P16] ③ Absence of phishing indicators in third party email and mobile client [P16]	3
Challenge 2. Content restrictions for phishing education and training E T	① Lack of engaging and interesting phishing education and training material [P10,P19,P28] ② Presence of complex interface and configuration in the game design [P28] ③ Repetitive training content [P7] ④ Disregard for user misunderstandings and interests [P11,P19] ⑤ Limited attack vector consideration [P19,P24,P59] ⑥ Disregard for both casual and serious gamers [P36] ⑦ Presence of cultural bias in the content [P36] ⑧ Time-consuming decision making process and lengthy training email [P5,P7]	9
Challenge 3. Design constraints for anti-phishing warning UI interfaces A	① Design similarity of phishing warnings with less serious security warnings [P1,P28] ② Frequent exposure causes warning fatigue [P4,P13,P14,P17,P18,P26] ③ Unsuitable warning placement [P2,P3,P5,P7,P11,P15,P25] ④ Absence of active user interruption [P1,P11,P14,P43,P44]	17
Challenge 4. Problems with anti-phishing warning content A	① Lack of comprehension and explainability [P14,P25,P49] ② Lengthy content [P41] ③ Distinct phishing warning design among vendors, platforms and web version [P49]	5
Ch5. Performance limitations of anti-phishing tools A	① Inadequate usability [P1,P2,P8] ② False positives and lack of reliability [P1,P2,P3,P8,P10,P13,P14,P18,P24,P25,P28,P44,P49,P57,P69]	15

Continued on next page

Challenges in anti-phishing intervention – continued from previous page

Challenges	Key points (included papers)	#
Challenge 6. Lack of attention to phishing indicators E T A	① Ignorance due to lack of trust and understanding on phishing warning and training [P1,P2,P3,P4,P8,P11,P14,P24,P28,P31,P36,P39,P44,P49] ② Disregard to warning due to appealing web content and site reputation [P2,P8,P14,P24,P49]	14
Challenge 7. Need to design specific training for spear phishing T	① Difficulty to detect spear phishing due to personal relevance and familiarity [P1,P7,P14,P15,P21,P26,P49,P58]	8
Challenge 8. Disregard for users' mental limitations during design E T A	① Users' distraction by other tasks is not well considered [P2,P7,P8,P13,P14,P24,P47] ② Users' inattentiveness to phishing interventions have not been taken into account [P7,P13,P14,P17,P24,P58] ③ Current design practices unconditionally rely on user decision [P4,P15,P17,P24,P25,P40,P49] ④ No alternative options for users to help them complete their primary task [P2]	14
Implementation		
Challenge 9. Anti-phishing technology deployment challenge E T A	① Deployment difficulty of anti-phishing technologies due to interdependancy on multiple factors and platform dependency [P23,P38,P50] ② Complicacy to safeguard employees in distributed and siloed settings due to enlarged attack surface [P6,P54,P57,P62,P65] ③ Training email spammed by email provider [P28]	9
Challenge 10. Technology adoption and usage challenges E T A	① Requirement of prior experience and investment in software for phishing games [P37,P45] ② Requirement of expertise and assistance from third-party services [P1,P8,P45] ③ Requirement of users' effort and willingness to use anti-phishing warnings [P19,P31,P45]	6
Challenge 11. Challenges due to complicated URL and domain name structures E T	① Similar organization name in the URL [P2,P45] ② Difficulties to detect minor changes in URLs [P46] ③ User confusion to identify phishing website hosted by trustworthy websites [P45] ④ Presence of textual manipulations and complex visual tricks in the URL [P45,P47]	4
Challenge 12. Obstacles to automate phishing incident response and anti-phishing training E T A	① Handling phishing incident reports requires the need for human validation [P50] ② Embedded training deployment requires manual human effort [P45]	2
Challenge 13. Exploitation of software vulnerabilities by attackers A	① Use of malicious javascript codes by attackers to bypass monitoring phishing plugins [P23] ② Use of XSS by the attackers to inject malicious code into legitimate webpages [P49]	2
Challenge 14. Unguarded email clients and websites A	① Limited use of SSL indicator to protect website login page [P2] ② No built-in mechanism in SMTP to prevent phishing [P16]	2

Continued on next page

Challenges in anti-phishing intervention – continued from previous page

Challenges	Key points (included papers)	#
Challenge 15. Limitations of current anti-phishing planning, policies and guidelines E T A	① Contradicting, incomplete and outdated anti-phishing recommendations in organizational websites [P15,P42] ② Choice of customized or outdated tools to manage IT incidents impact service quality and efficiency [P50] ③ Poor practice of training execution [P12,P59] ④ Lack of formal approach to gain experience from previous phishing incidents [P50] ⑤ Inadequate policies and guidelines to invoke user behavioral change [P38]	6
Evaluation		
Challenge 16. Lack of industrial relevance in evaluation practices and settings E T A	① The neglect of young people to test and improve their phishing knowledge [P21,P35] ② Sample bias due to limited demographic consideration [P1,P13,P14,P30] ③ Failure to conduct usability testing in real-world settings [P1,P7,P26] ④ Poor evaluation practices results in unreliable outcome [P14,P18,P32]	10
Challenge 17. Complications regarding data collection and replicating user experience E T A	① Difficulty to emulate users real-life experience in phishing studies [P3,P43,P31] ② Ethical difficulties of conducting phishing studies [P48] ③ Challenges of phishing study due to bias induced by the participants [P14,P21,P40]	7
Challenge 18. Insufficient usability and effectiveness evaluation of phishing interventions E T A	① Negligible practical value and effectiveness evaluation [P4,P8,P13,P18,P37,P40] ② Inadequate empirical investigation on variables used in phishing training and detection [P30,P41] ③ Lack of understanding on user behavioral response towards phishing incidents [P17,P33,P41]	10
Challenge 19. Lack of sophisticated quantification of phishing training outcome T	① Difficulty in measuring user phishing training effectiveness due to presence of bots [P55] ② Impact of pairie dogging on phishing training program outcome [P15,P59]	3
Challenge 20. Lack of post-training user knowledge retention practice E T	① Effectiveness of phishing interventions subject to dwindle over time [P13,P21,P40,P45] ② Lack of investigation on users' long term behavior change [P7,P31,P34,P54]	8

2.6 Critical Success Factors in the Design, Implementation and Evaluation

This section describes the data we collected to answer our second research question.

2.6.1 Critical success factors in the design

This section discusses factors documented in the literature that improve the design of current anti-phishing interventions.

CSF1. Design of engaging and up-to-date training content

Leveraging situated learning in anti-phishing training can improve user engagement. Adopting situated learning helps prepare users for a heavy cognitive load associated with experiencing a real-world phishing threat by introducing them to a relatable simulation scenario. Presenting information in an interesting way (e.g., Gamification, interactive training modules and videos, including less text, more graphics in the content, or comic format) brings enjoyment to the users, develops user confidence, strengthens motivation, and helps enhance content consumption [P5, P10, P19, P28, P34, P36, P37, P61, P62]. The training content should include different versions and varieties to allow individuals to learn in a way that suits them, as each individual learns differently. Also, phishing attacks are continuously evolving; training content can become outdated rapidly. It is crucial to include recent cyber attacks and detailed information about how attackers operate and the types of tactics used by attackers [P57, P59].

CSF2. Design of comprehensible anti-phishing technology

Identifying a phishing email or URL is a complicated task. Many basic concepts need to be explained to end-users to enable them to gain and apply the knowledge when they encounter an email with a suspicious URL. A comprehensive report during this process can help users make an informed decision and satisfy their curiosity when they seek more explanations to improve their contextual knowledge. A well-explained report aims to provide recommendations or feedback along with the reasons for the recommendations. An explanation that logically quantifies the decision made by the automated tool would increase users' trust in the systems. Users can also gain an understanding of the extent to which they can rely on the systems' decisions [P7, P8, P11, P14, P33, P39]. Offering anti-phishing recommendations to users in the form of visual examples and creating user-friendly URL patterns (e.g., using different colors for top-level domains and the rest of the URL) helps users to better absorb and retain the information [P8, P42].

In conventional phishing training or education methods, security experts decide what information to present to users. However, most often, security advice or help desk support is not available in real-time when the user experiences a phishing attack, during which there is little time to wait for expert advice or help desk support. Information from detailed reports can enhance the user's ability to comprehend the impacts of phishing attacks to provide appropriate responses to them in real-time. Providing an explainable report along with the automated anti-phishing technologies would encourage users to adhere to the warnings [P45].

CSF3. Diversity in training content to educate users on evolving phishing attacks

The diversity in learner behavior means that each user learns differently, therefore a cookie-cutter training method will only be effective for some users. It would be helpful to reach users in many different ways, for instance, by adopting various training methods such as flyers, posters, newsletters, and lunch and learn sessions [P58.] The continuously evolving nature of phishing attacks also demands changes and adjustments in the training content. A phishing game or training material containing only malicious domain and URL information leaves the user vulnerable to more advanced phishing attacks such as spear phishing. Leveraging diverse phishing attack vectors (link-based, data entry-based, or attachment-based) in the training template can enhance users' ability to detect a wider spectrum of diverse phishing attacks [P19, P61, P65].

CSF4. Consistency in training design

Consistency in training design helps users notice discrepancies, minimize confusion, and provide users with more opportunities to recognize inconsistent features. For example, a standardized template adopted by all the anti-phishing web pages reduces security fatigue. It also allows easier maintenance and updating. For instance, web designers can implement a tool more easily. To maintain consistency, researchers suggested a unified template for anti-phishing web pages proposed by a central agency such as CISA[6] or ENISA[7] [P42]. Studies also recommended that online services avoid using domain squatting techniques for domain names as it would be difficult for a user to identify the malicious domain if the legitimate domain uses domain squatting techniques, such as additional terms, unusual top-level domain (*facebook.me*), or subdomain (*extra.facebook.com*) [P46]. Using the same email styling in the organization is another recommended practice for design consistency [P41].

CSF5. Design of tailored phishing intervention

Poorly crafted and targeted phishing interventions will not be effective. For maximum reach and impact, phishing intervention should be appropriately personalized, for example, by including design of personalized training emails, adding local languages in the training content, designing realistic relevant template to train highly educated users, customizing the training style (where users can choose a preferred learning method), designing age-appropriate training tools (e.g., offering specific tools for children), dressing web application according to users' preferences, self-adaptive training where phishing simulation progresses in the level of difficulty based how well users perform, incorporating learning skills in the training design (e.g., consider casual and serious gamers in the phishing game design), customizing training content relevant to organizations and specific to job positions (e.g., managers or executives), and selecting training style suitable to the organizational settings (e.g., use of text-based

[6] https://www.cisa.gov/
[7] https://www.enisa.europa.eu/

training instead of comics) [P7, P16, P21, P26, P35, P36, P40, P48, P49, P52, P53, P57, P58, P59, P61, P62, P63, P64, P66, P67].

CSF6. Improving the UI design

Studies have suggested that UI designers should devote greater attention to designing effective UI for email clients and phishing interventions, in order to draw users' attention, remove user confusion, and better support user queries. For example, designers should focus on ensuring design consistency in the UI interface across mobile and web applications to help reduce mobile app users' exposure to phishing. Other design techniques include deploying a more visually salient interface with noticeable color variations, removing misleading UI phishing indicators for unverified emails, and adding a support icon in the email client and intervention UI design to support user investigation (e.g., adding a *"help me troubleshoot button"*) [P4, P5, P7, P16, P51]. End-users are encouraged to avoid using the same personalized indicators for different interfaces [P31].

CSF7. Design of informative and concise warning

Anti-phishing advice alone is not enough to modify user behavior and reduce their exposure to phishing attacks. Abstract information presented to users should be coupled with concrete examples to achieve more effective communication and information retention. Studies have shown that brief interventions have a relatively large positive impact. Too much information in the phishing intervention is unappealing, as inexperienced users may require an excessive amount of time to read and digest the information, creating an information overload. At the same time, interventions should be concise yet informative for educated and experienced users [P1, P5, P13, P18, P41]. Interventions should be designed in such a way that does not require lengthy decision making from users (e.g., hovering over a link in every email received) to save their valuable time. Studies have shown that many users do not click on explanatory buttons such as *"Lean more"* or *"More information"*. Therefore, warning designers should not hide critical information and should not require scrolling down or additional clicks before such information is revealed to users. If a *"Lean more"* button is deployed, it should contain very detailed information to satisfy user curiosity [P4, P5, P25]. When users encounter a warning, a clear choice or advice should be provided on how to proceed. Simply asking users not to proceed might be counterproductive. An alternative path can be provided to them to finish their task [P1, P2, P5, P14].

CSF8. Incorporating users' psychological and behavioral aspects in the design

A phishing attack takes advantage of users' cognitive limitations in order to succeed. Therefore, it is important to take into account the limitations in human cognition, user misconceptions (how attackers operate), user assumptions, decision-making process (e.g., what specific cues users look for and how users

interpret them), self-efficacy, and perceived threat in the design of phishing intervention [P9, P11, P18, P24]. Rather than informing users about the potential phishing risks, it is important to equip users to verify and assess the risks correctly. To design effective user-friendly phishing interventions, it is important to perform usability testing of phishing interventions to integrate user feedback in the design. Usability investigation allows the analysis of outcomes and helps identify trends over time to avoid repeating the same mistakes in the design [P22, P57, P61, P66, P67].

CSF9. Integrating phishing simulation with embedded training to facilitate education on demand

Organizations perform phishing simulations to test their employee's susceptibility to phishing. According to the existing PETA studies, coupling phishing simulation with training provides an effective approach to delivering an anti-phishing campaign. Studies have suggested that receiving immediate instruction after a user clicks on a phishing simulation link can guide users on diverse phishing tactics without being involved in an actual phishing attack. Accompanying phishing simulations with learning units helps achieve desired behavioral changes. Studies found that embedding phishing simulation with training increases the reporting of the actual phishing emails. Studies have shown that a substantial time lag between the cause (e.g., clicking on the phishing link) and the effect (e.g., getting a phishing warning message about the email) may confuse users about why they are receiving the subsequent warning message, as the time lag makes it more difficult for users to identify the original click which triggered the warning message. Instead of scaring and confusing users, the learning content encourages them to be more careful and attentive when they next encounter a phishing email. By offering training intervention to users straight away after they make a mistake, users will be more appreciative of the *education on demand* [P5, P7, P12, P27, P53, P57, P58, P59, P67, P68, P69].

CSF10. Focus on active warning designs

A warning should be designed without expecting the users to keep security in mind while they perform their regular online activities. Phishing concerns should be integrated into the critical path of users' primary tasks to force users to deal with any warning before proceeding, as this helps users shift their attention from their regular tasks to the phishing warning. Due to habituation, users are less likely to read phishing warnings in their entirety. Sometimes users ignore the passive warnings due to the design similarity with less serious warnings. Therefore, to increase user willingness to read the phishing warnings, the warning should be designed differently from other trivial warnings by employing design features such as varying text size, color, highlighting, distorting the visual appearance of a phishing website, and placing the warning close to the suspicious link [P1, P2, P14, P20, P22].

In phishing education and training, users are often asked to hover their cursors over a link to check the legitimacy of the link. Users may accidentally click on the suspicious link while hovering over it. To reduce the risk when users

hover over a link in a phishing email, researchers have suggested integrating action-based inhibitors (including a clickable link with a pruned URL) and adding a time delay before the link is clickable to allow users to overcome a small cognitive burden when dealing with the warning [P22, P25]. Adding interactive images on the login page, including *no-working links* on the site, is another possible forcing function that websites can adopt to prevent their users from submitting their credentials [P43, P44]. Adding an effective and non-obtrusive indicator on the site to notify users when they move from one domain to another (i.e., when clicking on a link leads them to an external website) provides another way of drawing user attention before they are exposed to potential suspicious URLs [P8].

2.6.2 Critical success factors in the implementation

This section describes the recommendations provided in the state-of-the-art research and industry practices to enhance the execution process of anti-phishing approaches, including anti-phishing interventions and technologies.

CSF11. Bringing key stakeholders on board to educate and encourage employees

In order to conduct a comprehensive and coordinated campaign of phishing simulation and training, C-suite executive officers are required to play a key role in its planning and implementation. For instance, extensive briefings before and during the phishing simulation and training are needed to better communicate with employees and resolve sensitive issues that may be encountered by individuals. Without such communication and sensitivity, a sudden and unexpected phishing simulation campaign may be criticized for undermining employees' self-esteem, targeting certain employees, or discriminating against specific groups. Managerial support from the C-suite helps employees understand that phishing security needs to be taken seriously. A group of champions must influence their peers about the necessity of phishing education and training and engage them in the process [P38, P40, P56, P57, P59, P61, P67, P68, P69]. University professionals and IT practitioners can also come forward to educate people about phishing. In this regard, some existing organizations such as $(ISC)^2$[8] help to empower professionals on every aspect of cyber security [P21]. Another good strategy is to leverage external service holders' capability to take advantage of their specialization. In this case, external vendors must understand particular organizations' cultures, requirements, and goals. Internal and external expertise can be blended to validate employees' behavioral change after knowledge-based assessment, develop phishing awareness materials, manage and track user engagement in phishing training, and fully manage every aspect of a phishing simulation and training program [P54,P60].

[8] https://www.isc2.org/

CSF12. Strengthen authentication and encryption mechanisms in browsers and email clients

Studies have recommended improving the existing authentication and encryption mechanisms in browsers and email clients to create a strong line of defense against phishing attacks. Examples include the use of an SSL indicator for protecting webpages, the use of a single domain name by companies to prevent users from becoming confused by multiple domain names or IP addresses, getting SSL certificates verified by a trusted CA [P2], deploying browser-based user authentication to draw user attention, improving server-side and end-user based security protection, adoption of SMTP security extensions (such as SPF, DMARC, and DKIM) by the email clients to authenticate incoming emails, use of security indicator to alert users when unverified emails reach user inbox [P8, P16], and developing anti-phishing tools or apps that can deactivate JavaScript when the focus is on an input field of a submission form and reactivate it when the focus is not on the form, in order to reduce the keystroke monitoring performed by the attackers to launch timing attack [P16, P22].

CSF13. Feedback, reminders, and reinforcement to maintain phishing awareness among users

Along with training and testing, research has shown that providing friendly reminders and helpful feedback throughout the intermediate micro-training facilitates positive behavioral changes among users. Reaching out to the users about intermediate results, early communication, and providing beneficial feedback can enable users to assess what they have learned and to improve on an individual level. Frequent notifications and reminders may result in information overload. Therefore it is recommended to send notifications only when security violations have occurred [P53, P58, P60, P61, P62, P69]. Rewarding users' positive behaviors, such as by providing certificates, positive reinforcement during group meetings, gift cards, or increments of time off, is also helpful in motivating users to act in security-conscious ways [P30, P61, P66].

CSF14. Conduct GDPR-compliant and anonymous training to protect user privacy and avoid false training outcome estimation

To be effective, phishing simulation emails are often enriched with users' personal information to train users in preparation for a spear phishing attack. Such personal information needs to be incorporated appropriately, should be GDPR compliant (for organizations that provide service to EU customers), and should be safe from a data protection perspective. Using straightforward phishing simulation emails, which are easy to recognize, might reduce user motivation by making users feel overconfident about their preparedness for such attacks. In contrast, employing phishing simulation emails that are too sophisticated might make users feel deceived and tricked. Therefore, a balanced mixture of easy and challenging phishing simulation emails should be used [P29, P69].

Keeping the simulation results anonymous and general, conducting phishing training as a whole rather than targeting individual employees, and formulating

the results in a way understandable to the users are among the recommended good practices [P59]. Finger-pointing of individuals who have failed a phishing simulation test can hinder their willingness to learn about phishing. For phishing simulations to be effective and sustainable, their primary emphasis should be on learning and maintaining anonymity (avoid collecting individual user's behavioral data). Individuals should not feel monitored and should be allowed to complete the phishing simulation at their own pace. Anonymous phishing simulation can also reduce the effect of prairie dogging. If prairie dogging occurs, even a zero-click rate after the phishing simulation would not be surprising. Prairie dogging distorts the phishing simulation outcomes and provides a false sense of an organization's preparedness for and susceptibility to phishing attacks [P59, P61, P62, P69].

CSF15. Providing phishing education and training to critical demographic groups

Research has shown that highly educated IT professionals are no less vulnerable to phishing attacks than non-experts. This indicates that phishing education and training are necessary for everyone. However, when there are resource limitations, training priority should be given to groups that are more vulnerable to phishing attempts (e.g., employees who have access to the shared network), less motivated, and more careless [P13, P40, P53, P58, P60].

Surveys with children and teenagers found that they need support to deal with phishing attacks. Parents must be well-experienced or sufficiently informed to teach their children about phishing. Therefore, it is recommended that anti-phishing education should be included in school curricula. This gives rise to another significant issue of educating educators. Incorporating phishing training and education into mainstream education requires training teachers to make them feel comfortable and confident to deliver lectures on phishing in their classrooms [P21, P35]. Phishing awareness campaigns should also consider educating retailers. Retailers should be trained to provide reliable and faithful trust signals in their website design. This will help customers distinguish between legitimate and fake websites and maintain the retailers' business reputation [P64].

CSF16. Automating the phishing training to support the organization's security teams

Automation of the creation and delivery of training content, as well as phishing incident response, management, and reporting, can assist IT security teams in several important ways, including saving resources, assisting organizations to stay on the right tracks, minimizing efforts in installation and maintenance, allowing rapid responses, limiting organizational damage, reducing the number of victims, and making the help desk perform more efficiently while maintaining good practices. Automation support at the help desk can accelerate critical assessments needed to determine whether phishing reports are genuine and consequently enable prompt responses to phishing attacks, especially given their complexity and scale. Full automation of phishing incident management and reporting might require much work to be achieved. Therefore, automation is

recommended for managing complex tasks when manual processes are too costly, which outweighs the cost of initial investment of installing such automated systems [P63, P67, P50].

Automation helps deliver personalized, frequent, and relevant training and helps with automatic threat identification and classification. Manually categorizing and personalizing the training content requires security teams to put in much effort and planning [P61].

CSF17. Better planning, policy management, and documentation on phishing training

Building sustainable phishing defense requires better policy-making and improved management. For instance, to enable simulated phishing emails to reach user mailboxes, IT systems require specific modifications. Otherwise, technical solutions deployed in the IT systems are likely to block the phishing simulation emails, assuming that they are harmful phishing emails. For example, these procedures include adding the email server's IP addresses to the technical IT system's whitelist and generating simulation emails that represent the entire spectrum of phishing simulations in order to test whether such phishing emails can manage to get through the automated phishing detection systems to reach the user mailbox [P69].

Organizations need to conduct market research in order to select a vendor to provide phishing simulation and training services, which best fit the organization's requirements. The research process includes analyzing reviews of the service provider from unbiased sources (e.g., G2[9]) to assess vendor popularity, getting advice from peers, and browsing vendors' websites to gain more detailed information about the services that they provide [P61].

While planning for a phishing simulation, companies should prepare their help desk to support user investigations [P51]. The planning process also includes communicating transparently with the employees and notifying them about the purpose of the simulations to prevent discomfort [P30, P69].

Effective phishing simulations need to be supported and facilitated by well-structured planning documentation about the phishing simulation and training. The documentation should follow the policy guidelines and standards that describe important terms (e.g., phishing, spear phishing, smishing, vishing, URL), cover all training types and contents, execution details, and frequency (e.g., how the training would be conducted, how many times training would be performed), expected behavior of participants, and rewards and consequences [P26, P60].

CSF18. Enabling and encouraging individuals to report phishing

Phishing reporting is important for defending organizations against phishing attacks and measuring the effectiveness of phishing simulation and training interventions. Reporting allows users to actively participate in phishing defense and help build a security-conscious culture. The reporting data also enables the organization's security team to analyze suspicious emails that have managed to

[9] https://www.g2.com/categories/security-awareness-training

bypass the technical phishing defense system, which in turn allows the security team to update and strengthen their automated phishing defenses accordingly [P58, P63]. Establishing a phishing reporting system before conducting the phishing simulation allows users to contact the IT security teams as soon as they encounter phishing. Integrating an easy-to-use and in-client reporting button can reduce the ticket volume and burden on the help desk, as simulated emails are not directed to the help desk. Users can be encouraged to report when they learn about the button [P26, P50, P58, P63, P69]. Training users on when and how to report and explaining to users the positive impact of reporting phishing can make a difference in establishing a solid line of defense against phishing attacks [P58].

CSF19. Invest in both technical and socio-organizational functions and capabilities

With the growing number of human applications and devices, it has become critical to address human risk factors. Recent investigations have demonstrated that cyber attackers have shifted their focus away from overcoming technology-based safeguards towards exploiting end-users instead. Technology-based solutions, software updates, security patches and firewalls sometimes fail to provide accurate detection (e.g., false negatives), leaving much of the responsibility of phishing detection to end-users. Therefore, successful phishing protection requires combining technology-based solutions with user-centric defense mechanisms. Coupling technology-based solutions with users-centric defense can reduce users' over-reliance on technical solutions [P3, P5, P12, P17, P26, P27, P28, P38, P41, P51, P53, P57, P58, P59].

To ensure better security and protection against phishing, a system that combines the strengths of two different detection approaches operating on different principles (e.g., combining blacklist- and whitelist-based phishing detection applications) can be deployed. This approach can take advantage of both applications, and one application could potentially detect a phishing email missed by the other [P18, P51].

2.6.3 Critical success factors in the evaluation

This section reports the suggestions collected from the primary studies on critical success factors that contribute to improving the evaluation of the effectiveness and usability of anti-phishing interventions.

CSF20. Conduct intermittent short-time progressive training to reinforce users' phishing awareness

The benefits of engaging in phishing simulation and training are subject to knowledge decay over time. Participants' knowledge about phishing after 6-8 months of training is similar to their pre-training level. Compliance frameworks like ISO 27001 and GDPR demand continuous employee training on cyber security topics, including social engineering attacks such as phishing, to develop a strong human-oriented line of defense. Due to this knowledge-waning effect,

to achieve desired behavioral changes, short training sessions, brief reminders, monthly phishing simulations with high-quality training materials delivered at the moment of failure (e.g., click on phishing link), and continuous training are recommended in the literature. Repetitive training is helpful for users who need help strengthening their full understanding, which may be only partially attained during the first time of training. Studies have shown that phishing click rates are significantly reduced after repetitive training sessions [P5, P7, P27, P34, P62, P67, P69]. However, maintaining an appropriate balance of training frequency is important to avoid training fatigue [P52]. With proper planning, phishing simulation emails should only be deployed occasionally while targeting a specific optimal frequency (e.g., 4-6 simulations per year). Excessive training will create an extra burden on the organization as this involves gathering and analyzing a large number of statistics and reports [P56]. Progressive training (easy-to-hard training) on more challenging topics would systematically improve users' sensitivity to deception cues [P24]. Sending emails in a randomized order and not flooding every department with phishing simulation emails would be beneficial, as this will reduce the chance of employees discussing the simulation with their peers and also minimize the ticket load experienced by help desks [P24, P53, P56, P57, P68, P69]. Evaluation of employees' knowledge retention immediately after training (short-term retention) and weeks or months after training (long-term retention) are essential steps in the continuous evaluation process to obtain the desired training effect [P52].

CSF21. Perform empirical testing and statistical analysis to improve and better support phishing training

Extensive empirical testing and evaluation of phishing simulation and training provide valuable data-driven insights, which can help organizations review their progress achieved by anti-phishing initiatives, optimize the training interventions, and plan long-term strategies to achieve more significant goals [P56, P57, P60, P61]. Setting a specific goal and establishing a baseline helps organizations stay focused on specific outcomes, recognize the specific changes required, track progress over time to ensure continuous improvement, and create a mature and robust phishing defense. Achieving cyber resilience takes time and requires patience and focused effort to assess users' shortcomings, in order to provide proper training [P54, P56, P58, P59, P60, P61, P68]. Existing studies recommend the establishment of a governance data structure for users to report to improve the capability to gather empirical data. Long-term impact assessment helps determine suitable training methods to increase user engagement [P31, P54, P57, P58]. Although regular, continuous training is recommended, regular training is costly and, in some cases, may be infeasible (e.g., training children in school regularly). Therefore, before conducting training, a rigorous evaluation of users' baseline knowledge should be performed. Moreover, challenging questions should be used in this initial evaluation to avoid the ceiling effect (minimizing the likelihood for participants to achieve the maximum test score in the initial evaluation). This will render future evaluation scores more informative as the increase or decrease of scores will be more visible [P21].

CSF22. Investigate if the phishing simulation is affected by false positives to avoid erroneous evaluation

Specific configurations (e.g., *presence of bots*) can cause false positives (e.g., high click rate) of phishing simulation assessment, leading to erroneous evaluation. Before conducting a phishing simulation, it is a good practice to check whether the current configurations will generate false positives. For example, teams responsible for phishing simulations should check the inventory of all software, security solutions, and service environments and documentation, in order to identify whether they are performing any scanning, probing, or analysis. If the answer is yes, it is vital to deactivate these capabilities for certain IP addresses to avoid generating false positives during the phishing simulation. Also, identifying the *bot clicks* by checking the clicks made by web browsers and operating systems. If an organization uses any email security solution with an allow-listing feature, it is recommended to prevent phishing links from being scanned or clicked by bots by identifying the *bot clicks*. To reduce false positives, participants should be informed to report phishing only through an approved reporting mechanism and avoid reporting phishing by using the email provider's default reporting button or function [P54]. Normalizing and re-scaling the click rates are also recommended to obtain a more accurate assessment of the outcome of phishing simulations [P32].

CSF23. Conduct user evaluation in their regular environment with realistic emails and measure delayed outcomes to replicate real-world settings

In real-world scenarios, users are not preoccupied with cyber security concerns when performing their regular online activities (e.g., checking email or conducting online shopping). Consequently, in a lab study, when users are asked to perform a security-related task (e.g., to identify phishing websites), users become more cautious, which disrupts their normal behaviors. Therefore, to produce generalizable results in a lab study, users' natural behaviors need to be preserved. While conducting phishing studies, researchers should observe participants' behaviors rather than interrupting; researchers should also obscure the purpose of the study by asking participants about other related subjects [P2, P4, P18, P43].

Evaluating user behaviors in their regular environment is an ideal approach, which can be achieved by employing the following recommended practices, to keep the experiment as realistic as possible and to achieve high ecological validity. Examples of these recommended practices include embedding simulation emails with users' regular email environments, asking users to install/deploy applications or browser extensions on their devices, collecting in-browser telemetry, and training users with realistic emails [P4, P7, P31]. A study has suggested collecting real-time neural and eye gaze data by *brain-eye* measure to evaluate the reliability of a user's response to a phishing study. According to the authors, if the neural features of the users show that they could have been more attentive during the study, their response might not be valid [P17].

Table 2.6 provides an outline of the critical success factors we assembled from the literature.

TABLE 2.6. Critical success factors in anti-phishing interventions

Critical success factors	Key points (included papers)	#
	Design	
CSF1. Design of engaging and up-to-date training content E T	• Incorporating situated learning to improve user engagement [P5,P10,P19,P28,P34,P36,P37,P61,P62] → Ch2.❶ • Including up-to-date content in the phishing training [P57,P59] → Ch2.❸	11
CSF2. Design of comprehensible anti-phishing technology E A	• Detailed report on anti-phishing efforts to persuade users to adhere to the warning and to support non-expert users [P33,P45] → Ch4.❹ • Explicit anti-phishing protection tools to increase users trust on automated anti-phishing tools [P11,P39] → Ch5.❷ • Integrate both visual and text example with explainability in the anti-phishing webpages [P42] • Designing user friendly URL bar to remove users domain name confusion [P8] → Ch11.❶,Ch.11❷ • Providing users with reliable automated anti-phishing tools [P7,P8,P14,P33] → Ch5.❷	8
CSF3. Diversity in training content to educate users on evolving phishing attack T	• Use of a variety of training content, a mix of tools for phishing training [P58] → Ch2.❸ • Attack vector variation in the phishing training content [P19,P61,P65] → Ch2.❺	4
CSF4. Consistency in the design E A	• Creating a standard unified template for anti-phishing webpages [P42] • Organizations should practice using the same structure and features for legitimate emails [P41] • Legitimate domain should avoid using common domain squatting techniques [P46]	3
CSF5. Design of tailored phishing intervention E T A	• Customised phishing training design for employees with power and authority in organization [P40] • Prioritising topics for training relevant to the organization [P16,P58] • Taking account the target demographic into training design and execution [P48] • Personalized training content [P26,P52,P53,P57,P59,P62,P66,P67] → Ch7.❶ • Considering casual and serious gamers need in the game design [P36] • Dynamic and self-adaptive phishing training [P63,P64,P66] • Personalized communication style and medium for phishing training [P61,P62] • Text training materials instead of comic materials in corporate settings [P7] • Developing anti-phishing tools for children [P21,P35] → Ch16.❶ • Web application dressing according to user preferences [P49]	21

Continued on next page

Critical success factors in anti-phishing interventions – continued from previous page

Critical success factors	Key points (included papers)	#
CSF6. Improving the UI design E A	• Disabling misleading UI elements for unverified emails [P16] → Ch1.❷ • Design of consistent phishing indicators for different interfaces [P16] → Ch1.❶ • Use of various colors [P5,P7] • Avoid using the same personalized indicators across different interfaces [P31] • Adding a support button in the email client to support user investigations [P51] • Adding an icon in email client indicating suspicious email [P7] → Ch1.❸ • Limiting the number of warnings user encounters to reduce warning fatigue [P4] → Ch3.❷	6
CSF7. Design of informative and concise warning E A	• Present abstract information using concrete examples [P1,P5,P13,P18,P41] → Ch4.❷ • Incorporate progressive disclosure in the design [P4,P5,P25] → Ch2.❸ • Warning should provide clear choice to the user [P1,P2,P5,P14]	9
CSF8. Incorporating users' psychological and behavioral aspect in the design T A	• Considering human vulnerabilities and decision making process in the design [P9,P11,P18,P24] → Ch8.❶,Ch8.❷,Ch8.❸ • Perform usability testing to improve warning design [P22,P57,P61,P66,P67] → Ch5.❶	9
CSF9. Integrating phishing simulation with embedded training to facilitate education on demand T	• Supplementing the phishing simulation with learning content [P5,P7,P12,P27,P53,P57,P58,P59,P67,P68,P69]	11
CSF10. Focus on active warning designs A	• Visual aids for safe browsing to draw user attention [P8] • Link focused warning in the email client to grab user attention [P25] • Warnings need to be actively interrupting users' primary tasks [P1,P2,P20,P22] → Ch3.❹ • Design of phishing warnings should be different than trivial warnings [P1,P14] → Ch3.❶ • Phishing indicators should distort the visual appearance of the website to help users distrust the phishing website [P1] • Warnings should stay long enough to grab users' attention [P1] • Action based inhibitor in the warning to reduce users cognitive burden and potential hazard of clicking malicious links [P22,P25] • Use of forcing and negative training functions [P43,P44]	9
Implementation		

Continued on next page

Critical success factors in anti-phishing interventions – continued from previous page

Critical success factors	**Key points (included papers)**	#
CSF11. Bringing key stakeholders on board to educate and encourage employees T A	• Important role should be played by the C-suite to secure the organization against phishing [P38,P40,P56,P57,P59,P61,P67,P68,P69] • Universities and practitioners should come forward to educate people [P21] • Leverage external service providers to support on phishing knowledge assessment and awareness material development [P54,P60]	12
CSF12. Strengthen authentication and encryption mechanisms in browsers and email clients A	• Use single domain name and use SSL to encrypt websites [P2] → Ch14.❶ • Deploying browser-based authentication [P8] • Adoption of SMTP security extensions in email applications [P16] → Ch14.❷ • Deactivate or re-activate javascript to avoid keystroke or timing attack [P16,P22] → Ch13.❶, Ch13.❷	4
CSF13. Feedback, reminders and reinforcement to maintain phishing awareness among users T	• Avoid frequent risk notification, avoid regular reminders, provide feedback to help maintain awareness [P53,P58,P60,P61,P62,P69] • Rewarding secure behavior [P30,P61,P66]	8
CSF14. Conduct GDPR compliant and anonymous training to protect user privacy and avoid false training outcome estimation T	• Conduct GDPR compliant phishing simulation [P26,P69] • Emphasizing the anonymity and learning aspect of the phishing simulation [P59,P69] → Ch17.❸ • Conduct random phishing simulation to reduce the effect of prairie dogging and estimate of organization's likelihood to fall victim to phishing [P61,P62] → Ch17.❸,Ch19.❷	5
CSF15. Providing phishing education and training to critical demographic group E T	• Raise retailers awareness about phishing along with their customers [P64] • Topics on anti-phishing training should be taught in the school to educate children [P21,P35] → Ch16.❶ • Everyone who has influence in organization's security should be trained [P53,P58,P60] • More focus on unmotivated and careless users [P40] • Teacher should be given priorities in terms of phishing education [P21,P35] → Ch16.❶ • Focus on vulnerable group for phishing education [P13]	8
CSF16. Automating the phishing training to support organization's security teams T	• Automation in delivering personalized contents and automation in threat identification [P61] → Ch12.❷ • Automating phishing reporting and incident response processes with the use of improved tools [P50,P63,P67] → Ch12.❶	4

Continued on next page

Critical success factors in anti-phishing interventions – continued from previous page

Critical success factors	Key points (included papers)	#
CSF17. Better planning, policy management, and documentation on phishing training E T	• Improved phishing defence through improved management and policy making [P11,P38,P40,P50,P53,P54,P57,P67] → Ch15.❺ • Structured and explainable policy and documentation of phishing training program [P26,P60] • Sending pre-notification to the participants to prevent discomfort [P30,P69] • Perform prior research and analyse the reviews on tools vendors [P61] → Ch15.❷ • Preparing IT system to avoid simulated email being filtered by technical filters [P69] → Ch9.❷ • Deploying post simulation help desk support to support users investigations [P51]	14
CSF18. Enabling and encouraging individuals to report phishing E T	• Establishing phishing reporting culture [P26,P50,P69] • Implementing easy-to-use, in-client phishing incident reporting tool [P58,P63] • Training users how to report phishing incident and explaining the benefits of reporting [P58,P60]	6
CSF19. Invest in both technical and socio-organizational functions and capabilities E T A	• Effective phishing detection requires the combination of technological innovation and human intervention [P3, P5, P12, P17,P26, P27, P28, P38, P41, P51,P53,P57,P58,P59] • Combining strengths of multiple anti-phishing technologies [P18, P51]	15
Evaluation		
CSF20. Conduct intermittent short time progressive training to re-inforce users' phishing awareness T	• Avoid over-training to reduce training fatigue [P52] • Multiple cycles of training to re-inforce phishing awareness [P24,P53,P56,P57,P68,P69] → Ch20.❶ • Repetitive training in a short time span [P5,P7,P27,P34,P62,P67,P69] → Ch20.❶ • Testing users' short-term and long-term knowledge retention after training [P52] → Ch20.❷ • Progressive training [P24]	13
CSF21. Perform empirical testing and statistical analysis to improve and better support phishing training T	• An extensive test with challenging question to reduce repetitive training cost and avoid ceiling effect [P21] • Conducting phishing simulation [P56,P57,P60,P61] • Assessment of long term impact [P31,P54,P57,P58] • Selection of effective metrics and relevant baselines [P54,P56,P58,P59,P60,P61,P68]	10
CSF22. Investigate if the phishing simulation is affected by false positives to avoid erroneous evaluation T	• Check if inventory management software are using any scanning, analysis or probing to identify unusually high volume of external IP addresses [P54] → Ch19.❶ • Normalize and re-scale click through rates for more accurate assessment [P32]	2

Continued on next page

Critical success factors in anti-phishing interventions – continued from previous page

Critical success factors	Key points (included papers)	#
CSF23. Conduct user evaluation in their regular environment with realistic emails and measure delayed outcome to replicate real world settings E T A	• Preserve users actual behavior to achieve results close to real world settings [P2,P18,P43] → Ch16.❸ • Use of field techniques for high ecological validity [P4] → Ch16.❸ • Testing users in their normal environment with instant corrective performance feedback [P7,P31] → Ch16.❸ • Realistic and equally difficult training emails to test the persistence of training outcome [P7] • Use of real-time brain-eye measure to collect transparent data [P17] → Ch17.❸	7

2.7 Insights from grey literature

The inclusion of practitioner perspectives and insights is critical for ensuring that phishing prevention strategies are effective and appropriately tailored to the evolving threat landscape. The grey literature in our study did not reveal any contradictory results compared to the academic literature. However, the grey literature was able to provide additional knowledge that would have remained undiscovered otherwise or reinforced and strengthened the results obtained from academic research. We further explain these points below.

• The omission of grey literature in our review can result in the loss of significant insights and information. Such a scenario may culminate in a disparity between the training administered and the practical threats that organizations confront, culminating in insufficient preparedness and heightened vulnerability to phishing attacks. For instance, important subject matters for inclusion in the design of phishing training content, as documented in CSF1 and CSF3, effective training techniques and strategies, as noted in CSF5, and recommendations for phishing training and incident response automation, as reported in CSF16, may be overlooked. Additionally, crucial aspects of training evaluation and knowledge assessment, along with suggestions for deploying and preparing IT systems for training, as highlighted in CSF17, might be excluded from consideration. Finally, valuable improvements to phishing reporting, as outlined in CSF18, may be missed, resulting in a lack of progress toward the enhancement of phishing defense mechanisms. Therefore, the inclusion of grey literature in our review is vital to ensure that all relevant knowledge is captured, thereby enabling the development of comprehensive and robust phishing education, training, and awareness interventions.

• Our academic research findings are substantiated and reinforced by data gathered from grey literature. For instance, an academic study (P52) revealed that over-training should be avoided to reduce training fatigue. However, several other academic studies (e.g., P24, P53) recommended repetitive training as an effective strategy for reinforcing users' phishing knowledge acquired during initial training. This proposition was supported by multiple grey studies (e.g.,

P56, P57, P68) documented in CSF20. Thus, the inclusion of complementary evidence from grey literature sources serves to bolster and augment the results obtained from academic research.

• The incorporation of findings from grey literature in our literature review has potentially broadened and deepened the scope of our analysis and strengthened the credibility and validity of our conclusions. For instance, several academic studies (P14, P18, P32) have assessed end-users phishing knowledge using click-through rates (e.g., the number of times users click on a phishing link), a practice we identified as problematic in challenge Ch15. The reliance on click-through rates as the sole evaluation metric can yield misleading and inaccurate results due to the presence of bots, as noted in the grey literature (Ch19). Accordingly, the grey literature emphasizes the importance of carefully selecting appropriate evaluation metrics and relevant baselines to avoid obtaining erroneous outcomes that could create a false sense of security within an organization (CSF21).

2.8 Limitations of this MLR

In this section, we discuss the internal and external biases induced in different stages of research methodology and the strategies undertaken to minimize them.

- As multiple researchers were involved in this study, to achieve *consistency* across different stages (study selection, data search, extraction), analysis and synthesis were performed according to a well-defined research protocol following a well-known MLR guideline [83]. Before finalizing our MLR research protocol, we modified and improved our MLR research protocol through a pilot study.

- A pilot study was conducted with 10 randomly selected studies. The corresponding data were extracted to ensure that our designed data extraction form covers all the relevant information (e.g., data related to our defined research questions). 90% (62/69) of the data was collected by the first author, and the second author extracted 10% (7/69). The data was shared in a shared folder and cross-checked by every author. Any disagreement among the authors was discussed and resolved in the weekly research meetings.

- Although *study selection bias* due to the impracticability of collecting a large number of primary studies is an unavoidable limitation in systematic reviews [112], [113], we endeavored to minimize the effect by systematically modifying our search string through a pilot study to capture all relevant academic studies.

- *Lack of generalizability* of the study outcome is another critical limitation common to all systematic reviews [113]. To ensure acceptable generalizability, we selected a popular digital library, Scopus, to collect our primary academic studies without restricting ourselves to publication year for high-quality venues. To collect our primary studies in the grey literature, we

chose Google as a search engine and iterated through numerous pages of search results until new pages of results no longer provided relevant information that related to the subject of our study.

- To minimize the *conclusion validity* bias originating from different interpretations of the exact result [113], the first author identified the codes and themes of the results, by applying an established method for analyzing and synthesizing qualitative data through thematic analysis, and shared the information with all other authors. Code books were updated based on the suggestions and feedback provided by other authors in weekly meetings.

- To ensure the *quality* of our grey studies, we executed a rigorous quality assessment by evaluating the grey studies based on five criteria (authority, methodology, date, novelty, outlet type) suggested by the MLR guideline [83]. Based on a pilot study, we define a minimum quality assessment score for the credibility analysis of studies in the grey literature.

- Another common threat to validity is *publication bias* arising due to researchers' tendency to report positive results compared to negative ones [82]. However, in our MLR, we reported adverse effects as challenges, thereby ameliorating potential publication bias.

Despite all the aforementioned mitigation strategies undertaken in this study, we acknowledge that our reported list of challenges and critical success factors may not be exhaustive due to the inevitable internal and external biases (e.g., missing primary studies, grey study quality assessment, usage of the non-comprehensive database, and thematic analysis process). Therefore, we encourage the readers to take this into consideration while reading our study as we believe that our study serves as a valuable starting point and one-stop-shop for readers to gain familiarity with the current state of practice in this domain, which can enable readers to explore possible areas that require further attention and investigation.

2.9 Discussion

In this section, we summarize and discuss the findings to provide an overall understanding of the key outcomes of our study. In relation to RQ1, drawing upon the evidence from 69 primary studies from both academic and grey literature, we discovered 8 design challenges, 7 implementation challenges, and 5 evaluation challenges. With regard to RQ2 (critical success factors), we identified 10 design CSFs, 9 implementation CSFs and 4 evaluation CSFs. We only summarize the highly reported challenges and critical success factors in the design, implementation, and evaluation stages of PETA.

- Our MLR uncovered the demand for improving the UI design of phishing warnings (a predominant design challenge discussed in 24% studies), specifically with warning design variation, active interruption, warning

placement, and warning exposure (Ch6). This evidence indicates the significant need to improve the design of the user interface of anti-phishing warnings to make them more accessible to end-users.

- A major challenge in the implementation of anti-phishing technology (Ch9) occurs during its deployment (13% studies), when problems arise due to platform dependency and distributed work settings. This suggests that developers and practitioners test their prototypes on different platforms before finalizing their models. This evidence also highlights the need to develop mechanisms that help organizations' security teams safeguard employees in a distributed office environment.

- Our findings disclose that the main evaluation challenges (discussed in 14% of studies) are limited industry relevance of the findings of studies (Ch16) and inadequate usability evaluation of phishing interventions (Ch18). This indicates the importance of *more rigorous* evaluation across a more significant array of demographic groups to test the usability of phishing interventions.

- In terms of design, the main critical success factor mentioned by 30% of studies is incorporating individual user needs into the design of phishing interventions (CSF5). This suggests the urgency of exploring the needs of individuals based on their age, educational qualifications, geographic location, profession, physical disabilities, language preferences, and other idiosyncrasies to improve their capabilities to detect phishing attacks.

- When it comes to implementation, most of our retrieved studies (21%) indicate that human-oriented education and training are of equal importance as adopting reliable technological anti-phishing solutions. Studies emphasized combining technical solutions that operate on different principles with providing education and training to users to reduce their dependency on technology-based solutions.

- Regarding evaluation, most of our studies (18%) recommended conducting follow-up training sessions to test users' knowledge retention and to reinforce users' phishing knowledge.

- In Table 2.6, we map the critical success factors with corresponding reported challenges. From this mapping, it is evident that challenge *Ch10. Technology adoption and usage challenges* do not currently have any recommended success factor documented in the literature that can help overcome the challenge. This provides an opportunity for future researchers to investigate such gaps in the literature, for example, by examining how the anti-phishing tools and applications can be simplified, how third-party dependency of anti-phishing tools can be minimized, how users can be encouraged to install and use the anti-phishing applications, and how requirements of user skills and experience can be minimized for application installation and usage.

2.10 Open Issues in Phishing Education, Training and Awareness Interventions

In the following sections, we discuss some open issues based on our findings and gaps in the literature identified in this MLR, which provide fruitful avenues of investigation for future researchers.

2.10.1 Equipping anti-phishing systems with explainable capability

In his study, Metaxas [114] states that *"It is the users' right and responsibility to decide what is acceptable for them. Their browser, their window to the cyber world, should enhance their ability to make this decision."* Our findings indicate that current anti-phishing tools are lacking in this respect, as they fall short of providing adequate explanations about specific phishing risks to users to enable users to make their informed assessments (Ch4❶). As a result, many users still fall prey to phishing even after receiving a warning due to a lack of understanding. This lack of understanding often leads to phishing warnings being ignored (Ch6❶). Several studies in our study pool recommended explainable and comprehensive anti-phishing tools to motivate users to adhere to phishing warnings by providing them with a detailed level of understanding of the reasoning (CSF2). Therefore an anti-phishing tool or mechanism should provide users with context-related information for various phishing-related problems that they may encounter. For example, assessing a phishing URL requires a user to have better knowledge about the structures of the URL to make an informed decision. To help users make an informed decision, browser developers can make the URL bar more user-friendly [P8]. Whenever an anti-phishing tool detects any unauthorized signal, it should provide users with adequate explanations and information about the warning, which would enhance users' knowledge about phishing and enable them to understand and assess future phishing risks [31]. A lack of information results in a lack of trust on the part of users and, consequently, users' under-reliance on anti-phishing tools [115]. In this regard, the system can offer additional information which helps users to make a correct choice [P33]. Explaining an anti-phishing tool's reliability (e.g., how the tool detects phishing attacks, its confidence level in its decision, and the logical consequences of a decision) also increases user trust and reliance on the anti-phishing tool [P33].

Practitioners who specialize in human-computer interactions can contribute to the design of anti-phishing tools that provide additional necessary information. When users encounter phishing attacks, they often seek suggestions from the help desk in making their decisions. Anti-phishing tools with analytical capability can reduce help desk traffic [P44]. Researchers and practitioners can borrow the concept of *explainability* [116] from machine learning and artificial intelligence in designing anti-phishing tools to provide additional information about the detection process applied by these tools. Research has shown that

explainability in the design helps users gain trust in the system [117]. Explainable anti-phishing tools can bring several advantages. For example, the logical reasoning about the underlying model would increase their justifiability and transparency, help debug certain flaws and improve the detection approach, and assist curious users in learning about the relationship and patterns employed to detect phishing attacks [118].

2.10.2 Platform for realistic phishing security testing

Our MLR reveals that existing phishing studies have faced difficulties in replicating users' real-life experiences with phishing within the study setting. Due to privacy reasons, it is difficult for the researchers to collect personal data, such as users browsing history, that could be useful in understanding user susceptibility to phishing and, accordingly, designing phishing interventions to educate and train users [P4]. Research outcomes can be affected by participants not following instructions (e.g., children discussing the study with their peers) [P21]; similarly, other studies documented that users did not follow or read the instructions before answering the survey questions [P17]. Some studies collected user feedback in an online survey or by adopting a role-playing scenario. Role-playing scenarios negatively impact users' security consciousness, as users tend to behave in less secure manners in the absence of real-world threats [P3]. Moreover, due to ethical restrictions, some studies reported the vulnerability to phishing by measuring only click rates. Ethical constraints also impose restrictions on how studies can be conducted. For example, users should not be tricked into giving away their personal information during the study [P16]. However, clicking is not the final stage in an actual phishing attack. Instead, the next step in a phishing attack involves users revealing personal information, which cannot be replicated ethically in a study setting. Game-based education typically trains users in an artificial environment which is dissimilar to the natural setting where phishing occurs [P26].

Some commercial phishing simulation and training platforms, such as knowB4 and Hoxhunt, provide organizations the opportunity to test users' phishing susceptibility in a relatively realistic setup. These platforms provide mechanisms to embed phishing simulation emails with clients' regular emails. Then a phishing training page will pop up if users click on the phishing link (to instruct users on how to respond when they see similar emails in real life). However, the effectiveness of this phishing simulation platform is affected by the impact of prairie dogging (challenge Ch19.❶), as bias is added when users know that their phishing knowledge will be tested [P21]. Therefore, to achieve the desired results, researchers and practitioners can focus on developing a platform that can provide a reliable mechanism close to users' real-life phishing experience to test phishing security. Researchers can take a conceptual idea from existing usability testing platforms such as Maze[10].

[10] https://maze.co/

2.10.3 Automated tool to assess users' attentiveness during online engagement

To improve the usability of the current phishing interventions, the literature suggests conducting usability evaluations and collecting users' feedback during their interactions with phishing interventions (e.g., CSF21). The design improvement primarily relies on users' feedback. Sometimes it is difficult to collect unbiased user feedback from surveys, as users try to hide negative experiences [P40]. Studies documented the need for a practical approach to identifying whether a participant has read the phishing intervention content [P7]. Designers of Phishing interventions can only improve their design if they receive reliable user feedback. We recommend that developers of anti-phishing tools should invest in the development of automated tools to identify users' attentiveness during online engagement, for example, eye-tracking and neuroimaging devices to collect users' neural and eye gaze features in real time to identify users' state of alertness. Although some studies have adopted a similar approach (e.g., P17), future investigation is needed to validate such a mechanism from a broader perspective.

2.10.4 Adopting automated and individualized approaches

Our MLR broadly highlights that a *cookie cutter* approach to phishing education, training, and awareness would not be adequate to cater to specific demographic groups across a wide spectrum. Aspects of phishing education and training, from training content to style [31]–[33], should be personalized to meet individual user needs to be helpful and effective. For example, a story-based training style is more effective for children [P21], whereas comic-based training content is unsuitable for corporate settings [P7]. Organizations should also consider modifying the content of their training program according to the routine business emails their employees receive. However, a fundamental challenge would be overcoming the manual effort required to personalize the training content. Incorporating automation can be a promising solution to conserve the time and effort of cyber security teams by replacing the manual process of choosing and modifying content suited to specific users. We suggest developing an adaptable phishing training approach where the difficulty level of the training content will be adjusted based on user knowledge. This can be achieved by conducting an initial phishing test to understand users' knowledge level and then delivering the training content automatically according to the users' skills and abilities. Automatically clustering users of similar knowledge can also help provide personalized training more efficiently. This will prevent individual users from becoming demotivated due to the training content not being pitched at their level of need.

2.11 Chapter Summary

This chapter answers the RQ1 and RQ2 of this thesis by providing a systematic overview of the socio-technical challenges and critical success factors in conducting PETA across the stages of design, implementation, and evaluation of these interventions. From 69 systematically selected primary studies, including 53 academic studies and 16 items of grey literature, we identified 20 challenges and 23 critical success factors. The identified challenges can help researchers and practitioners, particularly those new to designing, implementing, or evaluating anti-phishing interventions, to understand the ongoing potential obstacles and roadblocks that can hinder their endeavors from achieving anti-phishing effectiveness. Additionally, our findings on critical success factors reported in this chapter provide a foundation for the set of guidelines we present in Chapter 3 to support practitioners in improving the design, implementation, and evaluation of PETA interventions. Moreover, we provide a valuable mapping of the existing challenges to the critical success factors that can help mitigate each challenge. Finally, we highlight the remaining unresolved issues that require further attention and investigation, thereby providing a road map to inform future research.

Critical success factors to improve the design, implementation, and evaluation of PETA are piecemeal and scattered across a large number of sources, making it difficult for researchers and practitioners to understand how the success factors from different studies relate to one another. This limitation reduces the effective utilization of the existing findings in the prior literature to formulate a coherent set of guidelines to inform future practices. For example, two studies, P1 and P4, recommended improving anti-phishing browser warning tools; a study, P5, advocated for designing embedded phishing interventions. Yet another study, P6, recommended the design of user-friendly URLs, while several studies P11, P13, P35, and P36, recommended improving education related to phishing from a human-centric perspective. We systematically grouped these scattered recommendations from the existing literature to provide a list of logically organized and easily accessible critical success factors for researchers and practitioners.

This chapter provides *stage-specific* (design, implementation, and evaluation) challenges and critical success factors in PETA. Such stage-specific information is particularly useful in facilitating implementation in the industry, where different stages are performed by different stakeholder groups: specifically, the design of phishing interventions is typically conducted by design-oriented practitioners (such as warning designers), whereas the implementation and evaluation tasks are executed and managed by in-house cyber-security teams and/or senior executives of the organizations in question. This provides additional insights to enable us to devise *stakeholder-specific* guidelines targeting different practitioner groups discussed in Chapter 3.

Statement of Authorship

Title of Paper	Personalized guidelines for design, implementation, and evaluation of anti-phishing interventions
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in a manuscript style
Publication Details	O. Sarker, S. Haggag, A. Jayatilaka, and C. Liu, "Personalized guidelines for design, implementation, and evaluation of anti-phishing interventions," in Proceedings of the 17th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM), 2023, pp. 1–12.

Principal Author

Name of Principal Author (Candidate)	Orvila Sarker		
Contribution to the Paper	Performed analysis on all data, interpreted data, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	85%		
Signature		Date	20/12/2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Sherif Haggag		
Contribution to the Paper	Supervised development of work, helped in developing methodology, data interpretation and edit the manuscript.		
Signature		Date	20/12/2023

Name of Co-Author	Asangi Jayatilaka		
Contribution to the Paper	Supervised development of work, helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Please cut and paste additional co-author panels here as required.

Name of Co-Author	Chelsea Liu		
Contribution to the Paper	Supervised development of work, helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Chapter 3

Personalized Guidelines for Design, Implementation, and Evaluation of Phishing Interventions

Related Publication: This chapter is based on our paper: *"Personalized Guidelines for Design, Implementation, and Evaluation of Anti-phishing Interventions", published in the 17th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM).* (CORE ranking: rank A) [119]

Drawing insights from Chapter 2, it is evident from the literature that challenges hindering the success and effectiveness of phishing interventions need to be addressed. Our investigation reveals that the main reason that gives rise to the challenges can be attributed to the absence of personalized guidelines available to guide practitioners in addressing these socio-technical challenges.

To bridge this gap, in this chapter, we first systematically identify 22 influential factors—comprising 15 individual, 4 technical, and 3 organizational factors by analyzing the challenges and critical success factors identified in Chapter 2. Consequently, a collection of 41 guidelines is formulated to assist practitioners in tailoring the design, implementation, and evaluation of anti-phishing interventions. These guidelines are specifically designed for four distinct practitioner groups: designers/developers, information security teams, cyber security experts, and C-suite employees. Furthermore, the guidelines encompass 14 different sub-categories of interventions in phishing education, training, and awareness. It is important to highlight that our devised guidelines address 19 challenges out of the 20 challenges identified from the literature because only these 19 challenges apply to the practitioners involved in the design, implementation, and evaluation of phishing interventions.

The objective of these personalized guidelines is to enhance the effectiveness of current practices related to the development, deployment, and assessment of anti-phishing interventions. By disseminating these guidelines tailored to meet the needs of anti-phishing practitioners, we aim to contribute significantly to the ongoing endeavors aimed at mitigating the pervasive threat posed by phishing attacks.

3.1 Introduction

Empirical investigations have demonstrated that tailored design, implementation, and evaluation of phishing interventions can be efficacious in helping users recognize and mitigate phishing hazards [9], [25], [29], [30]. A phishing intervention refers to any anti-phishing system, software, tool, or framework that helps users deal with a phishing attack and requires user intervention [21]. The cognitive needs and mental status of individual end-users play an important role in determining the effectiveness of phishing intervention and, consequently, ought to be taken into consideration by the practitioners during the design, implementation, and evaluation process of phishing intervention [42]. However, evidence shows that practitioners often neglect end-users' decision-making processes and cognitive limitations in the design, implementation, and evaluation of phishing interventions. This leads to human-centric weaknesses or usability issues, rendering the end-users susceptible to phishing attacks [35]–[37]. To ensure the efficacy and usefulness of anti-phishing interventions, their key features such as the content and methods of anti-phishing intervention must be tailored to the needs of individual users [31]–[33].

The effectiveness of phishing education, training, and awareness (PETA) interventions significantly depends on decisions made by various practitioners involved in the design, implementation, and evaluation of these interventions. Studies document numerous examples of failures in such decision-making: some email providers do not use Simple Mail Transfer Protocol (SMTP) authentication mechanisms, thus allowing the attackers to send emails from spoofed email addresses [43]; many web developers of e-commerce enterprises fail to employ Secure Socket Layer (SSL) to secure their login page [36]; organizations' security officers often conduct phishing training without following any formal policies [47] and use unrealistic or irrelevant email templates [35], [48], [49]; developers design interventions with complex user interfaces that require specialized knowledge to install and utilize [23], [34], [50], [51]. These examples highlight the need for a structured set of guidelines provided to practitioners to help them comprehend the diverse range of needs and challenges end-users face. However, currently, there is a limited availability of such guidelines for practitioners. Moreover, most existing resources are intended for a singular group of practitioners or a particular class of cyber security interventions but are not specific to phishing interventions. For instance, Lujo et al. [56] and Lynsay et al. [57] reported guidelines for browser security warning design (this is a type of phishing intervention), and Mirium et al. [58] provided guidelines for the development of cyber security games (not directly related to phishing). Consequently, the target user group of these guidelines is primarily designers or developers of phishing interventions, with limited relevance to other practitioners such as organization managers or cyber security officers. Guidelines for IT security management (e.g., proposed by Pooya et al. [120] and Sonia et al. [121]) do not offer insights specific to phishing prevention. Overall, current guidelines and best practices for anti-phishing interventions (particularly on their design, implementation, and evaluation) are scattered around various academic and grey literature studies and not presented to practitioners in an

easily accessible and personalized format.

3.1.1 Research questions

In light of the needs and challenges faced by practitioners, we aim to investigate the following research questions in this chapter:

💬RQ3. What are the socio-technical factors that impact the effectiveness of phishing interventions in the design, implementation, and evaluation stages?

Motivation: This research question aims to discern the individual, technical, and organizational factors that play a role in either augmenting or hindering the overarching effectiveness of interventions to counteract phishing threats.

💬RQ4. What guidance can be provided to support the practitioners in addressing and incorporating the socio-technical challenges and factors in anti-phishing interventions?

Motivation: The motivation is to devise a set of guidelines by synthesizing the existing literature to aid practitioners in incorporating socio-technical challenges (discussed in Chapter 2) and soci-technical factors (outcome of RQ3) to enhance the effectiveness of the design, implementation, and evaluation of anti-phishing interventions.

3.1.2 Summary of the Findings

- By analyzing the challenges and critical success factors discussed in chapter 2, we have identified *22 dominant factors* which impact the effectiveness of anti-phishing interventions. These include 15 human factors (such as age, complacency, and educational qualification), 4 technical factors (such as device type and gamer type), and 3 organizational factors (such as organizational position and working hours). Our presented dominant factors can assist practitioners in attaining a deeper understanding of the important determinants of the success of phishing interventions.

- We have reported *41 guidelines* on the design, implementation, and evaluation of anti-phishing interventions, which are systematically compiled based on recommendations derived from the critical success factors discussed in Chapter 2, thus making the guidelines the first of its kind in terms of its comprehensiveness and breadth of coverage. Our devised guidelines can be a valuable resource to aid practitioners in improving the efficacy of anti-phishing interventions' design, implementation, and evaluation.

3.2 Methodology

Our methodology, summarized in Figure 3.1, consists of five steps, as detailed in this Section. It is noteworthy to mention that steps A (i) and the first part of A (ii) (thematic analysis of challenges) in Figure 3.1 represent the methodology

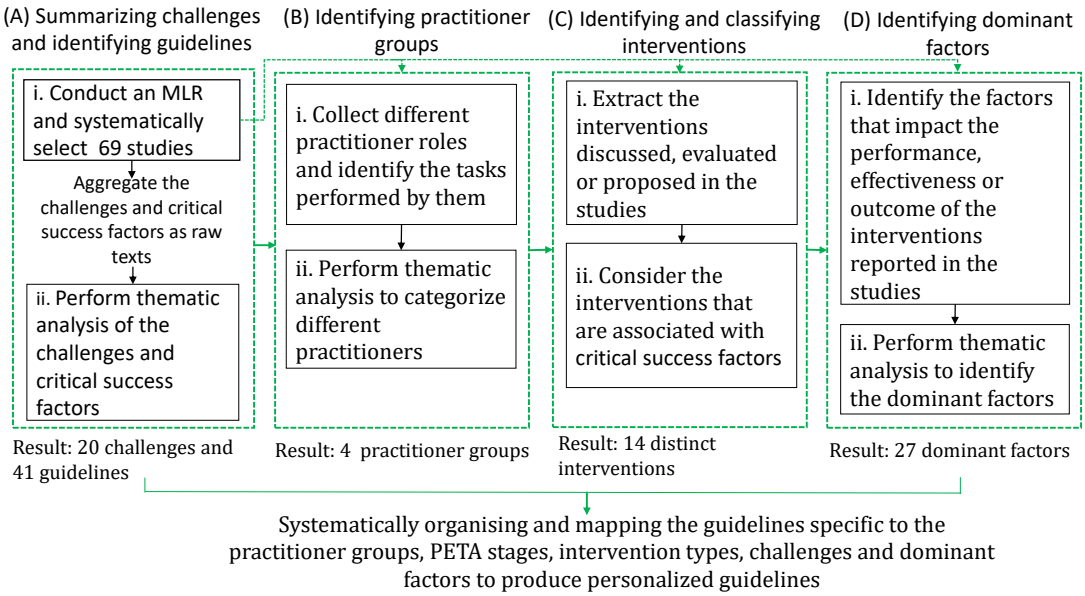

FIGURE 3.1. Methodology of this study

of our MLR discussed in Chapter 2. To assist the reader in better following the methodology employed in this chapter, we provide a short description of our methodology of MLR discussed in Chapter 2 in Section 3.2.1.

3.2.1 Summarizing challenges and identifying guidelines

Conduct an MLR and systematically select 69 studies

From the MLR discussed in Chapter 2, we identify the challenges and recommendations concerning anti-phishing interventions. As anti-phishing interventions are inherently industry-oriented, we incorporate the important perspectives of practitioners by including grey literature studies. Our MLR adheres to the protocols outlined by Kitchenham and Charters [82] and Garousi et al. (2019) [83].

In our study selection process, we utilized the Scopus [122] database, as it offers greater breadth and depth of academic literature compared to other databases [86], [87]. We ran a pilot study with ACM digital library and IEEE Xplore digital library to ensure that Scopus is comprehensive. Following a thorough examination of the search keywords used in prior review papers in this field (e.g., [16], [80]) and refinement through conducting pilot searches, we use the search keywords "aware*" or "interven*" or "nudge*" or "warn*" or "protect*" or "security indicators" or "alert*" along with the keyword "phish*" to collect the academic studies. To ensure the quality of the collected studies, we only included studies that had a CORE [123] rank of A*, A, and B. The CORE ranking system aims to ensure high-quality standards and rigorous peer review processes for selected journals and conferences [124]. We omitted papers with a CORE rank B published before 2012. The reason for this is that our pilot study revealed that several CORE B papers published before 2012 proposed client-side anti-phishing tools without conducting a real-world evaluation of their usability

(e.g., Passpet [125]). We excluded studies that have provided automated anti-phishing solutions (defined in [77]) and research that were not written in English, short papers (less than 6 pages), and literature survey papers.

We employed Google as our search database to collect studies in the grey literature. Google serves as a widely accepted and utilized search engine for gathering grey literature study [88], [90], [91]. For the collection of studies in the grey literature, we used the search terms "education", "training", "awareness" with the term "phish*". The rationale for employing an alternative search query distinct from the one employed in the academic study stems from the search methodology employed by Google; Google conducts searches using the specified search terms across its entire index of webpages [92]. Hence, apart from the duplicated webpages and academic studies already identified in Scopus, our pilot study using the identical search string employed in academic studies yielded numerous extraneous results. Therefore, we used different search terms to obtain more relevant results following the suggestions by Garousi et. al [83].

The grey literature study collection process concluded on Google's 16th page as no new or redundant information was identified, as recommended by Garousi et al [83]. To ensure the quality of the grey literature studies, we assessed the publication's authority, methodology, presence of reliable references, date of publication, the novelty of the article, and the article's outlet type as suggested by Garousi et al [83].

After applying the inclusion/exclusion criteria and data quality assessment, we collected a total of 69 studies, including 53 academic and 16 grey literature studies. We use the symbol P[*] to denote the studies throughout the rest of the chapter.

Perform thematic analysis of the challenges and recommendations

To identify the challenges and recommendations documented in the literature, we utilized thematic analysis - a standard data analysis method for qualitative data - to process the raw textual data (challenges and recommendations). In particular, we adhered to the process outlined by Braun and Clarke [106] by using Nvivo - a tool designed for qualitative data analysis [106]. After extracting the textual data into an Excel spreadsheet, we imported the data into Nvivo to perform open coding, which involves breaking down the data into smaller components and assigning labels to each component [107]. This process was conducted iteratively, with codes generated in the initial stage being modified and updated in later stages. Examples of the data analysis process for challenge and guideline identification are shown in Fig. 3.2 and Fig. 3.3 respectively.

From the analyzed data, we identified 20 challenges in the design, implementation, and evaluation phases of phishing education, training, and awareness interventions. We classify the challenges into three broad categories: first, design challenges relate to the concept, functionality, feature, or user interface of anti-phishing interventions. Examples of design challenges include *inconsistent UI design across browsers and mobile devices* [P22, P49], *unsuitable warning placement* [P2, P3, P5, P7, P11, P15, P25], and *complex interface and configuration in the game design* [P28]. Second, implementation challenges relate to intervention automation, deployment, and adoption (e.g., *the interdependency*

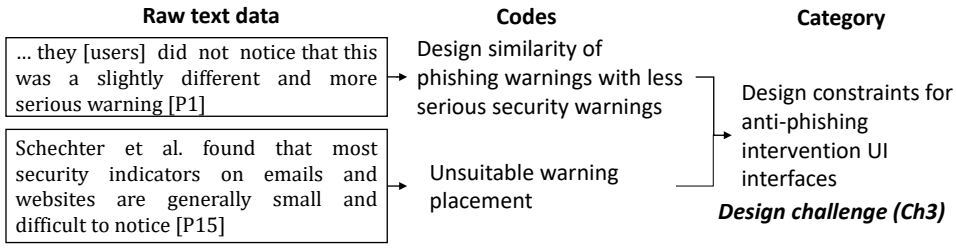

FIGURE 3.2. Identification of the challenges using thematic analysis

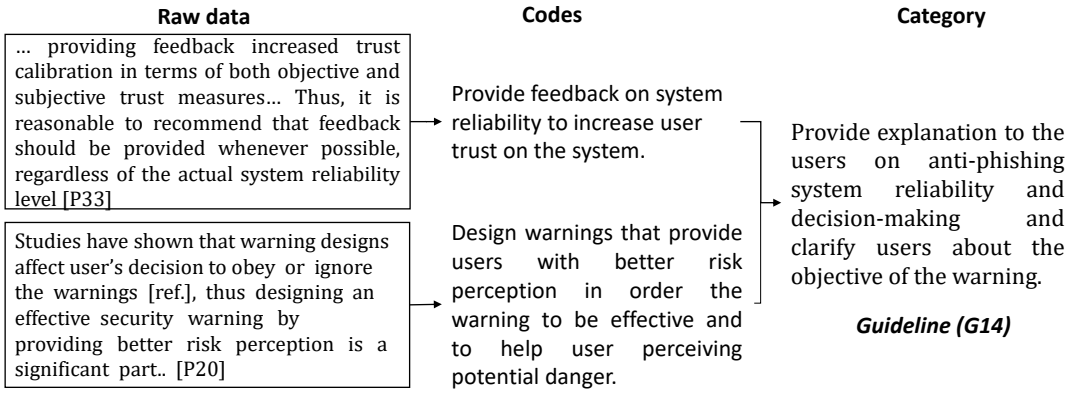

FIGURE 3.3. Identification of guidelines using thematic analysis

between different factors and platforms in implementing effective anti-phishing measures [P23, P38, P50]). Third, evaluation challenges relate to measuring the usability and effectiveness of anti-phishing interventions and quantifying training outcomes. As practitioners are the primarily targeted user groups of our personalized guidelines, we excluded challenges relevant exclusively to researchers (i.e., Challenge 17 - limited demographic consideration in the study settings discussed in Section 2.5.3).

We collected several critical success factors documented in the literature, such as "feedback should be provided whenever possible, regardless of the actual system reliability level [P33]", "designing an effective security warning by providing better risk perception is a significant part [P20]". We performed the thematic analysis of these success factors to report 41 guidelines.

3.2.2 Identifying practitioner groups

To report personalized guidelines for anti-phishing intervention practitioners, it is important to identify their specific needs and responsibilities involved in anti-phishing interventions' design, implementation, and evaluation processes. To derive a mapping between practitioner groups and their roles and responsibilities, following existing studies in other domains (e.g., [126]–[128]), we categorize different practitioner groups based on their functional roles as documented in the literature. From the literature, we first gathered information on various practitioner roles, precisely 28 different groups, such as browser developer [P3, P8], platform designer [P4], designers of anti-phishing applications [P18], game developer [P36], information security officer [P27], chief information security

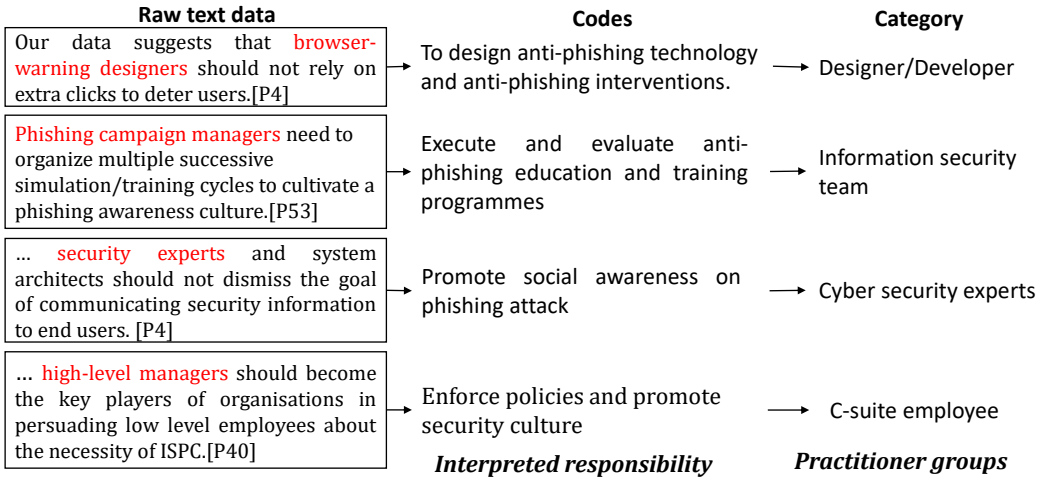

FIGURE 3.4. Categorization of different practitioner groups using thematic analysis

manager [P34], and cyber security practitioners and decision makers [P53]. We then carefully investigated the responsibilities performed by each practitioner group based on the critical success factors collected from Chapter 2. For example, from the following textual data "*phishing campaign managers need to organize multiple successive simulation/training cycles to cultivate a phishing awareness culture [P53]*", we infer that one of the responsibilities of *phishing campaign managers* is to execute and evaluate anti-phishing education and training programs. Fig. 3.4 shows an example of how we identified the responsibilities of different practitioner groups from raw textual data in the literature.

Based on these identified responsibilities of each practitioner group, we then categorized all practitioners involved in anti-phishing interventions into four major categories, namely *designer/developer*, *information security team*, *cyber security experts*, and *C-suite employees*. Table 3.2 summarizes the responsibilities of different practitioner groups. Some of the responsibilities overlap across different practitioner groups. For example, a designer/developer whose main responsibility is to *design* anti-phishing tools may also *evaluates* an anti-phishing software [P22, P57, P61, P66, P67] because a designer/developer may need to test the usability of interventions before finalizing the design. However, as shown in Table 3.2, organizations' information security teams mainly carry out evaluating anti-phishing interventions.

3.2.3 Identifying and classifying interventions

For each study collected in our MLR, we carefully examined the introduction, methodology, and results sections in search of any interventions that were suggested, debated, or evaluated. Consequently, to tailor the guidelines to each type of intervention, we only included those treatments that had pertinent recommendations reported in the research. We classified the interventions into three types - education, training, or awareness - based on characteristics such as their intended goal, presentation, and method of delivery of anti-phishing information to users (the terms phishing education, training, and awareness

are defined in Table 1.1). We identified two types of phishing education: *anti-phishing instructions* and *educational games*. Identified phishing training interventions are: *phishing simulation and embedded training*, *phishing training game*, *narrative-based training*, *instructor-based training*, *information and guidance-based training*. For phishing awareness interventions, we found *email client phishing indicators*, *browser SSL warnings*, *browser EV certificate warning*, *browser security toolbar*, *browser phishing warning*, *QR code scanner phishing warnings* and *Interactive custom phishing indicator*. Table 3.1 displays the definition of each subcategory of interventions within phishing education, training, and awareness.

TABLE 3.1. Identified intervention types

No	Intervention	Description
		Phishing education
E1	Anti-phishing instruction	Anti-phishing instructions (e.g., anti-phishing webpages [P42], posters, or leaflets on how to deal with phishing [P13]) aim to guide the users about several aspects of a phishing attack, for instance, about the best practices to recognize phishing attacks, instructions on how to report phishing attacks.
E2	Educational game	Taking a didactic approach, phishing educational games provide users with information and resources to help them better understand the topic of phishing [P9, P10, P11, P36, P37].
		Phishing training
T1	Phishing simulation and embedded training	In phishing simulation and embedded training, organizations send simulated phishing emails to their employees from a specialized phishing simulation software or service to test employees' vulnerability to phishing attacks. In most cases, after an employee takes a harmful action (e.g., clicks on a suspicious link, provides sensitive information, etc.), they are presented with training or learning content that explains the consequences of their actions [P5, P7, P12, P26, P27, P32, P38, P40, P53, P55 to P69].
T2	Phishing training game	Phishing training games often adopt a hands-on, experiential approach to train users about phishing [P19, P28].
T3	Narrative-based training	In narrative-based training, users are provided information in the form of stories [P15, P48].
T4	Instructor-based training	In instructor-based training, users are delivered tutorials on phishing by a security expert (e.g., chief information security manager [P34], a training expert [P21]).

Continued on next page

Identified intervention types for guidelines - continued from previous page

No	Intervention	Description
T5	Information and guidance-based training	In this form of training, users are explained certain facts about phishing and then they are advised some guidelines on what to do when they encounter a phishing attack. Unlike instructor-based training, this type of training can be provided by anyone, for example, a security expert or a peer [P15, P24].
	Phishing awareness	
A1	Email client phishing indicator	Email client phishing indicators work by scanning incoming emails and analyzing various factors such as the sender's email address, email content, links and attachments or the email header to identify potential phishing attempts [P16, P25].
A2	Browser SSL warning	A Browser SSL warning appears when there is a problem with the SSL certificate for a website. It verifies that the website being accessed is legitimate and that the connection between the user's browser and the website is encrypted [P4, P8, P14, P22].
A3	Browser EV certificate Warning	A browser EV (Extended Validation) certificate warning, on the other hand, provides a higher level of security and validation. In addition to verifying the domain, an EV certificate requires additional verification of the organization or entity behind the website, including legal and operational checks [P8].
A4	Browser security toolbar	Browser security toolbars are add-ons or extensions that can be installed on a web browser to provide additional protection against phishing attacks. These toolbars may include features such as URL scanning, page analysis, real-time updates, and user reporting to help detect and block known phishing websites. Additionally, some browser security toolbars may include other security features, such as blocking pop-ups or disabling scripts, to further protect the user against malicious content [P2, P18].
A5	Browser phishing warning	Browser phishing warnings (active or passive) are built-in security features of web browsers [P1, P4, P14, P17, P23, P25].
A6	QR code scanner phishing warnings	QR code scanner phishing warnings provide anti-phishing warnings to users by checking the links contained within the scanned QR code against a database of known malicious URLs [P20].
A7	Interactive custom phishing indicator	Interactive custom indicators force the user to interact with her customized (personal) indicator (image/text) to log in [P3, P43, P44].

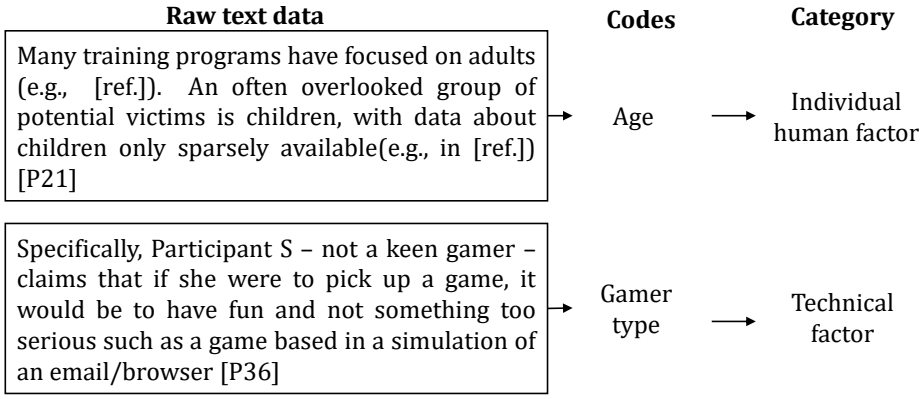

FIGURE 3.5. Dominant factor identification from raw text data

3.2.4 Identifying dominant factors

From the raw texts collected on the challenges and critical success factors relating to anti-phishing interventions (as discussed in Section 3.2.1), we investigated the main *reasons* for poor outcomes of existing anti-phishing interventions, as well as areas of improvement suggested by the authors to achieve better user experience. Based on this synthesized information, we identify and coin the term *dominant factors*, which refers to the individual, technical or organizational factors that may either enhance or impede the overall outcome of anti-phishing interventions. These factors are called *dominant* as these factors were argued to influence the outcome of the anti-phishing interventions after empirical evaluation with the users in the results and discussions of our collected studies. We adopt the terminology utilized by prior researchers [16], [129], [130] to designate the dominant factors identified in our investigation.

As an example of identifying a dominant factor, from the textual data *"staff may expect to learn more from experts while [college] students may expect to learn more from their peers [P48]"*, we derived an understanding that, according to the authors, the education qualifications of users have a significant bearing in determining their preference for the type of training methods and their effectiveness. This information indicates that designers could consider the educational qualifications of the end-users to improve user experience with the phishing intervention. Accordingly, the dominant factor identified here is users' *educational qualification*. Fig. 3.5 illustrates an example of dominant factor identification from the raw textual data.

Fig. 3.6 depicts an example of the interconnection among challenges, guidelines, practitioner groups, interventions, and dominant factors derived from the raw text data. To determine the interconnection between challenges and interventions, we identify the interventions discussed in the study and the limitations mentioned within those interventions. Similarly, in establishing the interconnection among guidelines, interventions, and practitioner groups, we search for critical success factors that are directed toward practitioners to enhance the outcomes of specific interventions.

TABLE 3.2. Identified practitioner groups and their responsibilities

No	Practitioners	Responsibilites/Activities
U1	Designer/ Developer	• Design and deploy anti-phishing technology and anti-phishing interventions [P4, P19, P54, P60, P61, P65]. • Stay informed and up to date about the latest tactics and techniques used by cyber attackers to initiate phishing attacks to update the design [P19, P61, P65]. • Collaborate with other security professionals to ensure that anti-phishing interventions address all critical aspects of cyber security [P54, P60]. • Conduct usability testing of anti-phishing intervention to improve the design [P22, P57, P61, P66, P67]. • Provide support to the client organization's security team to deploy, execute, and evaluate anti-phishing technology and anti-phishing interventions [P54, P60].
U2	Information security team	• Implement the anti-phishing technology and anti-phishing solutions within the organization [P26, P50, P53 P61, P63, P67]. • Execute and evaluate anti-phishing education and training programmes [P30, P54, P59, P61, P62, P69]. • Prepare and manage organization's IT system to run anti-phishing training and handle phishing incidents [P54, P69]. • Identify the vulnerable employees within the organization who require education and training [P26].
U3	Cyber security experts	• Make decisions on the appropriate elements and aspects to be included in the anti-phishing interventions [P13, P41, P42], promote social awareness on phishing attack [P4, P21]. • Responsible for what should be given priority in terms of educating and training end users [P41]. • Responsible for promoting social awareness of phishing attacks [P21].
U4	C-suite employee	• Enforce policies to educate and train employees against phishing attacks [P11, P38, P40 P50, P53, P54, P57, P67] and to promote security culture [P21, P26, P38, P40, P50, P56, P57, P59, P60, P61, P67, P68, P69] within the organization. • Collaborate with the organization's security team to adopt strong anti-phishing measures and prepare and execute a phishing incident response plan [P53, P56, P57, P60, P61]. • Motivate and encourage employees to act securely [P21, P38, P40, P56, P57, P59, P61, P67, P68, P69].

3.3 Human-centric and Socio-technical Factors Impacting Anti-phishing Interventions

We analyzed the human factors discussed by Dupont [129] and refined by Desolda et al. [16] in the context of phishing attacks. A total of 8 human-oriented factors identified from our textual data (complacency, distraction, lack of communication, pressure, lack of knowledge, lack of resources, fatigue, and norms)

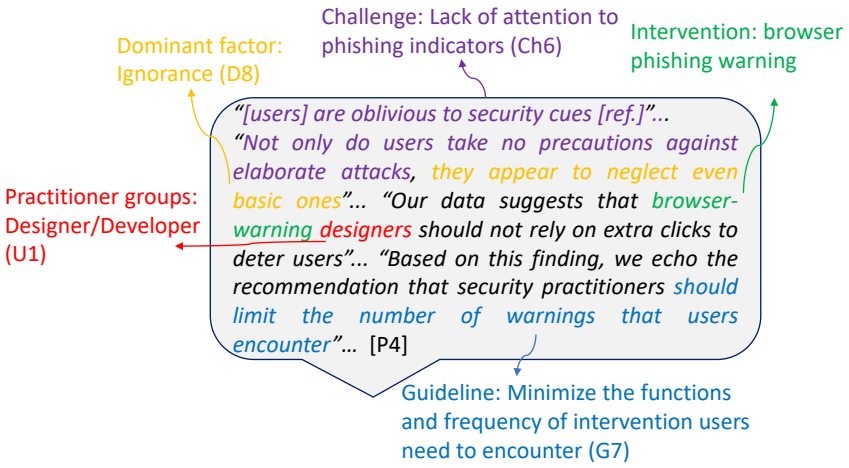

FIGURE 3.6. Example interconnection among challenges, guidelines, practitioner groups, interventions and dominant factors

confirm the factors documented by Dupont. A number of additional factors have emerged from our textual data, prompting us to search the literature for examples of their manifestation in order to categorize these factors. Following the approach by a study in the field of software engineering by Dulaji et al. [130], we were able to identify and name the other previously unexplored factors, which make a significant contribution of new knowledge to this field.

We identified 22 dominant factors to underscore the significance of incorporating user needs and preferences in the design, implementation, and evaluation of anti-phishing interventions. These dominant factors emphasize that neglecting user's requirements and inclinations can hinder practitioners' efforts in providing tailored design, implementation, and assessment of a system, which in turn results in compromised usability and suboptimal outcomes. Table 3.3 details our 22 identified dominant factors and how their inclusion or absence can influence the outcome of anti-phishing interventions. We grouped our dominant factors into three categories: individual human factors, technical factors, and organizational factors, as detailed in the following sections.

3.3.1 Individual human factors

Based on the analysis of selected studies, we identified that various demographic characteristics of individual users require greater attention from practitioners to enhance the efficacy of anti-phishing interventions across different user groups. According to the literature, practitioners need to take into consideration a number of user demographic characteristics, including *age* and *educational qualification*, in order to enhance the efficacy of anti-phishing interventions across different user groups.

In addition to demographic characteristics, our additional dominant factors also relate to cognitive functioning and limitations of individual users in

order to provide a comprehensive picture of user-level characteristics. Specifically, *knowledge decay*, *distraction*, *lack of attention* and *lack of motivation* all constitute individual limitations that can reduce user's capacity to effectively identify or counteract phishing attacks. For example, studies have shown that the knowledge acquired by individuals through phishing training tends to degrade or dissipate over time [P7, P13, P21, P31, P34]. Again, phishing warnings frequently pass unnoticed as users are preoccupied with concurrent tasks or are incapable of maintaining attention across multiple stimuli [P13, P14, P41].

Our investigation reveals that certain personality traits or user characteristics, such as *complacency* and *optimism bias*, may lead to users disregarding anti-phishing interventions. For example, users tend to overestimate the efficacy of their organization's anti-phishing measures [P7] and exhibit over-reliance on website content that is visually attractive [P2, P5, P8, P11, P49]. It is noteworthy to mention that our dominant factors represent distinct attitudes or mindsets. For example, complacency involves a belief that things are good enough as they are and that there is no need for further effort or change [131] which may cause complacent individuals to overlook potential phishing risks. Conversely, optimistic individuals feel less likely to experience cyber attacks which may cause them to share passwords, visit untrustworthy websites and so on [132].

An absence of tailored design, implementation, and evaluation of anti-phishing interventions by taking into consideration user needs and preferences could result in overwhelming the user and causing *security fatigue*. For example, fatigue can result from frequent exposure to warning and risk notifications, thereby reducing their effectiveness [P4, P13, P14, P17, P18, P26]. Similarly, excessive training can lead to training fatigue [P34, P52, P53, P58, P60, P61, P62, P69]. The complexity inherent in software installation procedures, as well as the intricacy of the process for reporting phishing incidents, may lead to *lack of user motivation*.

3.3.2 Technical factors

The effectiveness of anti-phishing warnings is greatly influenced by the *type of devices* utilized by the users.

Table 3.3. Dominant factors and their impact on anti-phishing interventions

No	Factors	Impact	#
		Individual human factors	
D1	Age	• Children aged 8-13 require specialized phishing educational intervention as they are biologically less attentive [P21, P35]. • Teenagers tend to make decisions quickly without considering the consequences, and are more susceptible to being persuaded by urgency and panic-inducing phishing emails [P35]. • Older employees have relatively bad training outcomes as they prioritize maintenance over growth [P40]. • Age 18-25 are vulnerable to phishing attacks [P6].	4

Continued on next page

Dominant factors and their impact on PETA – continued from previous page

No	Factors	Key points (included papers)	#
D2	Complacency	• Users' overconfidence in the appealing web content leads them to disregard phishing warnings [P2, P5, P8, P11, P49]. • Users' prior experience with websites results in overconfidence, causing them to disregard phishing warnings [P1, P2, P5, P11, P13, P25, P44]. • Users over-rely on site reputation and trust the warning [P14]. • Users are overconfident about their ability to detect phishing [P43, P44, P45], over-trust on their organizational technical phishing solutions [P7].	14
D3	Confusion	• User confusion arises due to similarity in domain names [P45], webhosting [P45, P49], distinct warning design patterns among vendors [P25, P49] and conflicting information present in the anti-phishing guidelines [P42]. • Users become confused about the purpose of a received training email [P5].	5
D4	Curiosity	• Users click on the phishing link out of curiosity [P2, P25].	2
D5	Distraction	• Users are distracted by other tasks as security is not their main concern [P13, P14, P41]. • Individuals are unable to focus on multiple things simultaneously (e.g., noticing on phishing warning while doing online shopping) [P13].	3
D6	Educational Qualification	• Phishing stories from a peer is an effective method of training for college students [P48]. • University staffs learn better from facts from an expert-based training method [P48]. • Compared to bachelor's degree, users having master's and PhD degrees are more confident in detecting phishing [P52].	2
D7	Knowledge decay	• The knowledge gained by users during phishing training tends to dissipate over time [P7, P13, P21, P31, P34].	5
D8	Ignorance	• Users failed to look at anti-phishing interventions [P7, P13, P17, P28], ignored as web content looked legitimate [P2] and when received a high frequency of warnings [P4].	6
D9	Lack of communication	• Before designing and implementing anti-phishing software, users' interests and needs are not well investigated [P16].	1
D10	Lack of motivation	• Users are not motivated enough to install anti-phishing software on their devices [P31, P37], show unwillingness to report phishing due to a complicated reporting process [P50, P58, P63] and do not find the training and educational material interesting [P10, P19, P28].	8
D11	Lack of trust	• Users do not trust anti-phishing warnings due to limited accuracy of anti-phishing tools [P1, P11].	2
D12	Optimism bias	• Optimistic users tend to be less conscious as they believe that negative events only happen to others [P13].	1
D13	Perceived vulnerability and severity	• An individual's heightened understanding of the consequences of phishing attacks enhances their resistance to these types of attacks [P40].	1
D14	Pressure	• Phishing incident response by IT staff gets delayed due to the reception of a high volume of phishing reports [P50]. • An individual receiving a high volume of emails is more susceptible to phishing attacks [P26].	2

Continued on next page

Dominant factors and their impact on PETA – continued from previous page

No	Factors	Key points (included papers)	#
D15	Fatigue	• Providing comprehensive instruction could result in overwhelming the user [P13]. • Frequent exposure to warning causes warning fatigue [P4, P13, P14, P17, P18, P26]. • Frequent risk notifications and excessive training result in training fatigue [P34, P53, P58, P60, P61, P62, P69].	13
		Technical factors	
D16	Device type	• Individuals who rely on mobile devices are at a higher risk, as phishing signs are obscured or not fully visible on the small screens of mobile devices [P49].	1
D17	Gamer type	• A casual player is unsatisfied with playing a phishing game that is designed for serious gamers, and conversely, a serious gamer is unfulfilled playing a phishing game that is intended for casual players [P36].	1
D18	Lack of knowledge	• Users do not understand anti-phishing warnings due to lack of knowledge about security and security indicators [P1, P4, P5, P6, P7, P8, P9, P10, P11, P13, P14, P17, P20, P21, P28, P35, P39, P46, P47, P49].	20
D19	Lack of resource	• User do not have enough infrastructure support when they work from home [P6]. • Absence of abstractness in the anti-phishing recommendations and lack of advanced anti-phishing tools reduces users' self-efficacy [P42]. • Users do not receive training emails due to emails being in the spam folder [P28].	3
		Organizational factors	
D20	Organizational position	• Employees in a higher position in an organization are more vulnerable regardless of the phishing training or punishment [P40].	1
D21	Social influence	• People trust others' phishing stories as they perceive this information as trustworthy [P15]. • Observing others share information results in heightened levels of disclosure [P13]. • The motivation, self-efficacy, and cognitive ability of employees are impacted by the social relationships within and surrounding the organization [P26, P40].	4
D22	Norms	• Organization's procedural measures (e.g., security policies, standards and guidelines) have a beneficial effect on raising security consciousness [P38].	1

Research has shown that web developers tend to avoid adding phishing indicators to mobile browsers to save space for web content [P16]. Based on a research report by [P36], it is evident that anti-phishing educational games fail to tailor their content to the specific interests of individual user groups. To elaborate, the games designed for serious gamers do not meet the expectations of casual players, and vice versa. Consequently, designers of anti-phishing games ought to take into account the *type of gamers* as a factor when creating an educational game that can cater to the unique requirements of both casual and serious gamers.

According to multiple studies, the challenge faced by users with limited technical knowledge is attributable to their insufficient familiarity with security indicators and tools, as well as the complexity of the requirements of third-party

tools [P1, P4, P5, P6, P7, P8, P9, P10, P11, P13, P14, P17, P20, P21, P28, P35, P39, P46, P47, P49]. This highlights the significance of incorporating the needs of novice users into the design process. Additionally, the effectiveness of users in detecting and preventing phishing attacks is reduced by a lack of resources, such as infrastructure support and advanced anti-phishing tools, and the absence of abstractness in anti-phishing recommendations [P6, P42].

3.3.3 Organizational factors

We identified several organizational factors from the literature that could be incorporated to enhance the effectiveness of the anti-phishing interventions. For example, according to research, *higher-positioned* employees in an organization are more susceptible to phishing attacks, regardless of their previous training or negative experience of being victims of phishing attacks [P40].

The literature shows that the implementation of security policies, standards, and guidelines in an organization is beneficial to increasing phishing awareness among employees. Additionally, the positive impact of organizational *norms* and *normative beliefs* on employees is observed in their exercise of greater caution, when opening potentially harmful phishing emails [P38]. Employees' motivation, self-efficacy, and cognitive ability are affected by social relationships [P26, P40]. For example, according to several studies [P13, P15], individuals tend to trust and share phishing stories based on the perceived trustworthiness of the information and the influence of social relationships within and around their organization.

3.4 Personalised Guidelines for the Design, Implementation and Evaluation of Anti-phishing Interventions

We present our guidelines and the rationale underpinning each guideline in Table 3.4. In the following paragraphs, we briefly discuss the guidelines personalized to different practitioner groups, intervention stages, intervention types, challenges and dominant factors in anti-phishing interventions.

• Among the 41 guidelines, 20 are mapped as relevant to our first user group (U1) consisting of designers and developers. This user group has the highest number of guideline relevant to them. These guidelines (G1 to G11, G13, G14, G16, G17, G19, G21 to G23, G27) include recommendations on interface design, placement of the phishing indicator, intervention content design, user engagement strategies, attention-drawing techniques, and enhancements for existing and future intervention designs.

• Our second user group (U2) consisting of *information security teams* of the organizations have a total of 19 guidelines (G12, G15, G16, G18, G20, G24, G25, G28, G29, G31 to G38, G40, G41) mapped to them as potentially applicable. These guidelines are intended to assist these cyber security professionals in conducting effective phishing education and training sessions, implementing measures to reinforce the organization's security, enhancing the speed and

efficiency of phishing incident response and reporting procedures, as well as improving the educational resources made available to users to better cope with the threat of phishing.

• The guidelines G25 and G26 have been devised for *cyber security experts* (U3)), with a primary focus on conducting an investigation on the anti-phishing solution before adoption, as well as implementing protocols to ensure organizational adherence to a standardised email and anti-phishing webpage template.

• The guidelines G26, G27, G30, and G39 have been reported specifically for *C-suite employees* (U4) within organizations. While these guidelines are primarily intended for use by C-suite employees, certain guidelines, such as G28 and G29, are also relevant to the organization's security team (U2). These guidelines emphasize the importance of collaboration between the C-suite and the security team to develop policies that meet the needs of employees and provide better support for victims of phishing attacks.

• The user group U1 is mapped to guidelines for all three stages of phishing intervention, whereas the user group U2 are linked to guidelines solely for the implementation and evaluation stages. In contrast, user groups U3 (cyber security expert) and U4 (C-suite employee) have guidelines exclusively for the implementation stage of phishing intervention.

• Across the three stages of anti-phishing interventions, we provide 19 guidelines (G1 to G11, G13, G14, G16, G17, G19, G21, G22, and G27) for the design stage, 17 guidelines (G12, G15, G16, G18, G20, G23 to G28, G30 to G32, G35 to G37, and G39) for the implementation, and 8 guidelines (G13, G16, G29 to G33) for the evaluation of anti-phishing interventions. It is noteworthy that some guidelines are applicable to multiple different stages. For example, G16 is applicable to the design, implementation, and evaluation stages. We arrived at this recommendation based on several studies (e.g., P7, P13, P15, and P16) that suggest incorporating users' preferences in the design (e.g., the layout of anti-phishing intervention), implementation (e.g., training methods used to educate users), and evaluation (e.g., email templates used to assess users' phishing knowledge) of phishing interventions.

• The guidelines presented in this study have the potential to address several challenges in the design, implementation, and evaluation of existing anti-phishing interventions. For example, these guidelines can be particularly useful in addressing design constraints specific to anti-phishing warning user interface (e.g., G7, G8, G9), improving content-related issues for phishing education and training (e.g., G4, G5, G6), mitigating issues related to the deployment and adoption of anti-phishing technologies (e.g., G18 and G7), overcoming limitations associated with existing anti-phishing planning, policies, and guidelines (e.g., G12, G24, G25), enhancing the efficacy of current evaluation practices and addressing challenges associated with evaluation settings (e.g., G16, G30, G31), and retaining user knowledge after training sessions (e.g, G27, G28, G29).

TABLE 3.4. Guidelines for anti-phishing interventions

No	Guidelines	Rationale
G1	Remove deceptive user interface elements for unverified emails and incorporate an alert icon within the email client to indicate potentially fraudulent emails.	• Disabling misleading UI elements (e.g., profile photo, email history) for unverified sender addresses will reduce user confusion [P16]. • Placing a security indicator for unverified email delivered to the user acts as a forcing function for the sender domain to configure their SPF/DMRC/DKIM correctly [P7, P16].
G2	Clearly display the underlying URL of a suspicious link in the email client	• Clearly displaying the underlying URL of a suspicious link in the email client (link-focused warning) make it easier for users to notice where the links' actual destination [P25].
G3	Incorporate progressive disclosure in the design and add a learn more button.	• Progressive design and learn more buttons help to facilitate general advice, satisfy user curiosity, and support user investigations [P4, P5, P25, P51].
G4	Use visual examples and explanations and avoid technical jargon in the content.	• Avoiding technical details in the content can make them understandable to non-expert users [P1]. • Integrating visual examples and explanations on phishing cues presented helps users memorize and understand better [P42].
G5	Present abstract information and leverage situated learning in the content.	• Leveraging situated learning in the design can make the intervention interesting and engaging, and also improves learning outcomes [P5, P10, P19, P28, P34, P36, P37, P61, P62]. • Too much information in the content can be unappealing to inexperienced users [P1, P5, P13, P18, P41]. • Adopting situated learning is beneficial as learning science suggest that simply asking users to follow some advice would not be helpful [P5].
G6	Introduce varieties in the content and keep the information up to date.	• Including varieties in the content can help users tackle new and emerging phishing attacks [P19, P57, P58, P59, P61, P65].
G7	Minimize the functions and frequency of intervention users need to encounter.	• Limiting the frequency of the warnings reduce warning fatigue [P4]. • Minimum number of functionalities in the game can help finish the game easily, easy for users to remember when functionalities are less [P10].
G8	Design phishing warnings differently from standard warnings.	• Variation in the design increases the likelihood for users to read it, ensures they are taken seriously and prevent habituation [P1, P2, P14].
G9	Make the critical information easily accessible and visible to the users.	• To make users easily notice the warnings [P1, P4, P8, P25], increase warning adherence [P25] and to impose forced attention [P8, P25].
G10	Create uniform phishing indicators across different browsers and mobile interfaces.	• This will reduce the susceptibility of mobile device users [P16].
G11	Provide users clear choices and actionable items to proceed.	• Active interruption and actionable items minimize the user's workload, are naturally noticeable and users can use their time efficiently [P1, P2, P4, P5, P7, P20, P22, P24, P25 P41, P43, P44]

Continued on next page

Guidelines for anti-phishing interventions – continued from previous page

No	Guidelines	Rationale
G12	Offer intervention immediately after users fall for phishing.	• Avoiding delay in displaying warnings minimizes users' confusion [P5]. The right timing of training intervention provides instant education [P2].
G13	Perform usability tests and collect user feedback.	• Collecting users' feedback from usability testing can improve future intervention design [P18, P22, P57, P61, P66, P67].
G14	Provide an explanation to the users on anti-phishing system reliability and decision-making and clarify users about the objective of the intervention.	• Feedback on the anti-phishing system increases users' trust [P7, P8, P11, P14, P33, P39, P43], helps users perceive potential danger [P20], increases user understanding and improves user ability to detect phishing [P18, P39]. • Making it clear to the users why they have displayed the intervention or not taken to the website to avoid their confusion [P5,P14].
G15	Use both technical and human-centric defence mechanisms to cope with phishing.	• Prevent user's over-reliance on technology, provide additional defence in detecting unpredictable, highly dynamic, and increasingly sophisticated phishing attacks [P3, P5, P12, P17, P18, P26, P27, P28, P38, P41, P51, P53, P57, P58, P59]. • Educating users about the security properties of different interventions remove their misunderstanding that leads to mistake [P14]. • Training all individual who has access to the organization increase the organization's robustness [P53]. • Human-centric defence mechanisms organized by C-suit employees can help low-level employees in the organization to learn about phishing [P21, P38, P40, P56, P57, P59, P61, P67, P68, P69].
G16	Personalize the intervention style and medium based on the target user's demographic.	• Personalized phishing training can take into account user's preferences (e.g., individual preferred training method [P15, P21], content relevant to the organization [P16, P58], roles and responsibilities [P40, P53, P58, P60], age [P21, P35]) to ensure users receive targeted education and training [P7, P13, P15, P16, P21, P26, P35, P36, P40, P48, P52, P53, P57, P58, P59, P60, P61, P62, P64, P66, P67].
G17	Consider the decision-making process and vulnerabilities of humans in the design.	• Taking into account the vulnerabilities and decision-making processes of the user (e.g., users' misconceptions and perspectives [P11], perceived threat [P9]) increases the effectiveness of anti-phishing interventions for end users and assist to develop the tailored approach [P4, P6, P7, P9, P11, P18, P24].
G18	Configure IT system for phishing training.	• Preparing IT system to avoid simulated email being filtered by technical filters helps users being missed for training [P69]. • Verifying if inventory management software is utilizing scanning, analysis, or probing techniques help detect abnormally high levels of external IP addresses [P54].
G19	Design visually distinct user-friendly URL bar.	• Noticeable and consistent URL bar helps users differentiate legitimate and malicious domains easily [P2, P8, P46].

Continued on next page

Guidelines for anti-phishing interventions – continued from previous page

No	Guidelines	Rationale
G20	Use automated platforms and improved tools for phishing training, incident management and reporting.	• Automated approaches help to better support managing complex situations, delivering personalized content and threat identification [P61, P63, P67, P50].
G21	Disable JavaScript on login forms when a form element is in focus.	• Deactivating JavaScript on webpages every time the focus is put on a form element prevents the attacker from capturing the keystrokes or initiating timing attacks [P16, P22, P23].
G22	Explain the capabilities and effectiveness of the deployed anti-phishing solution clearly to the users.	• Reliable trust signals to the users can prevent over-trust and over-reliance on the deployed anti-phishing solutions [P11]. • Utilizing interactive error messages to elucidate the purpose of a website can deter users from engaging in destructive actions [P43, P44].
G23	Use email authentication protocols to encrypt emails and filter out incoming malicious emails.	• To achieve better resiliency [P18,P51] and to make more informed decision [P16, P27] on the incoming emails.
G24	Send pre-notification to the users before conducting phishing training, however, perform random phishing training.	• Sending pre-notification to the participants prevents discomfort [P30, P69]. • Emphasising on the anonymity of phishing training can reduce the effect of prairie dogging and estimate of organization's likelihood to fall victim to phishing [P59, P61, P62, P69].
G25	Conduct prior investigation before adopting anti-phishing tools, identify most vulnerable group and determine priority topics.	• Perform prior research and analyze the reviews on tool vendors to select the right tool [P26, P61] • Identifying vulnerable users can help reduce training time and efforts [P26]. • Teaching everything or huge amount of information can cause security fatigue [P13].
G26	Follow a consistent template for organizational emails and create a standard template for anti-phishing webpages.	• A consistent email structure helps employees to notice the discrepancies in phishing emails easily [P41]. • A standardized template for anti-phishing webpages reduces inconsistency helps avoid confusion and helps web-designer implement their anti-phishing tools easily [P42].
G27	Introduce a user-friendly, built-in phishing reporting tool within the client system. Develop a formal procedure to handle phishing reports.	• Having a formal procedure placed makes it convenient to handle phishing reports [P50]. • An in-client phishing incident reporting tool makes phishing reporting easier [P58, P63].
G28	Get employees' feedback to modify the organization's policy.	• Obtain staff's feedback after phishing simulation to modify the organization policy accordingly to meet staff's needs [P50].
G29	Deploy help-desk and victim support for users.	• Deploying post simulation help desk support allows further users' investigations [P51]. • Deploying help-desk support can assist external users in determining the authenticity of an email sent from the organization [P51]. • Add a victim support option in the anti-phishing webpages can help users to fix potential problems [P42].

Continued on next page

Guidelines for anti-phishing interventions – continued from previous page

No	Guidelines	Rationale
G30	Create a structured policy and documentation. Regularly assess and manage phishing awareness efforts.	• Appropriate policy and documentation ensure that all the employees adapt themselves to security countermeasures and requirements [P26, P38, P60]. • Continuous measurement, improved management and policy making helps to achieve improved phishing defence [P11, P38, P40, P50, P53, P54, P57, P67].
G31	Conduct phishing simulation with embedded training.	• Assist the organization's security team in practicing the handling and response to simulated phishing incidents to enhance preparedness for real phishing attacks [P53, P56, P57, P60, P61]. • Embedding learning content with phishing simulation provides education on demand [P5, P7, P12, P27, P53, P56, P57, P58, P59, P60, P61, P67, P68, P69].
G32	Conduct phishing simulation that adheres to the guidelines of the data privacy policy appropriate to the region.	• Data privacy policy-compliant phishing training protects participants sensitive information, hence reducing data breaches [P26, P69].
G33	Provide users immediate feedback on their performance.	• Users feel motivated if instant corrective feedback is provided after testing and evaluating their phishing knowledge in their regular environment [P7, P10, P31].
G34	Use realistic and equally difficult training emails. Use challenging questions to test phishing knowledge.	• Realistic and equally difficult email helps to test the persistence of the training's effect [P7]. • An extensive test with challenging questions reduce repetitive training costs and can help avoid the ceiling effect [P21].
G35	Implement progressive and self-adaptive phishing training.	• Dynamic and self-adaptive phishing training improve user sensitivity to deception cues [P24, P63, P64, P66].
G36	Adopt video and interactive education and training materials.	• Video and interactive training are more effective as users do not need refreshment very quickly [P5, P11, P19, P34, P36]
G37	Utilize the expertise of external service providers to aid in phishing knowledge assessment and awareness material development.	• Leveraging external service providers can support better phishing knowledge assessment and awareness material development [P54, P60].
G38	Choose evaluation metrics and baselines that are useful and relevant.	• Click-through rate should be normalized based on the persuasiveness of the training template to produce a sound analysis and evaluation [P32, P54, P56, P58, P59, P60, P61, P68].
G39	Train users how to report phishing and reward secure behaviour.	• Training users on how to report phishing incidents and explaining the benefits of reporting can help to establish a phishing reporting culture [P26, P50, P58, P60, P69]. • Rewarding employees for their secure behaviour can motivate and encourage them to perform better [P30, P61, P66].

Continued on next page

Guidelines for anti-phishing interventions – continued from previous page

No	Guidelines	Rationale
G40	Conduct multiple cycles of follow-up training.	• Help to assess users' short-term and long-term knowledge retention after training [P26, P31, P52, P54, P57, P58]. • Repetitive training in a short period helps users learn a second time if they had difficulty understanding in the first time [P5, P7, P24, P27, P34, P53, P56, P57, P62, P67, P68, P69]. • Follow-up training (for children) to counter knowledge decay of the ability to identify phishing [P21].
G41	Avoid frequent reminders and over-training and keep the reminders short and simple.	• Avoiding frequent risk notifications and over-training reminders can reduce training fatigue [P34, P52, P53, P58, P60, P61, P62, P69]. • Including a lower bound of information in the reminder measures can reduce security fatigue [P34].

• Among our 3 main intervention types (i.e., education, training and awareness), our analysis has yielded a set of 9 guidelines for education interventions (G5, G7, G16, G17, G25, G26, G29, G33, G36), 27 guidelines for training interventions (G5, G6, G7, G11 to G16, G18, G20, G24, G25, G27, G28, G30, G31 to G39), and 19 guidelines for awareness interventions (G1 to G5, G7 to G11, G13 to G17, G19, G21 to G23).

• Some of our challenges (i.e., "Challenge 5 - performance limitations of anti-phishing tools", "Challenge 18 - insufficient usability and effectiveness evaluation of phishing interventions", "Challenge 19 - lack of sophisticated quantification of phishing training outcome") do not exhibit any discernible dominant factors. As a result, the guidelines G13 and G38, which are intended to address these challenges, are not linked to any particular dominant factor.

3.5 Limitations of this Study

One limitation of our study arises from the fact that design, implementation and evaluation principles are not always explicitly articulated [77], consequently, we cannot claim that our proposed guidelines in this study encompass an *exhaustive* list of all potentially relevant principles. The presented guidelines have been formulated with specific regard to the collected study context being investigated, and therefore, their *generalizability* may be limited. As a means of substantiating these standpoints, we posit that the 69 studies used in this research were collected through a rigorous process of quality assessment. Most of the guidelines that we have formulated are supported by more than one study in the literature, which underscores their applicability in contexts that are different from the collected study context. This is because these studies have involved diverse user types, varying sample sizes, different intervention types, and other variables. Additionally, our analysis encompassed industry reports [e.g., P58] and case studies with various organizations [e.g., P63, P64, P65]. The inclusion of grey studies facilitates the mitigation of bias that stems from a proclivity to publish studies that report favorable results exclusively [133].

Regarding the *representativeness* of the data used in this study, the collection of textual data is restricted to 53 academic and 16 grey studies which have been identified in literature searches. We recognize that future research could broaden the scope of the searches and analysis by including additional data sources such as interviews and surveys [134]. In order to validate and strengthen the usability and effectiveness of our guidelines, we plan to conduct a semi-structured interview study with developers and cyber security practitioners in our future work.

As multiple researchers were involved in this study, in order to minimise *researcher bias*, various activities (e.g., study selection, data search, extraction, analysis, and synthesis) were conducted in accordance with a well-defined research protocol, following the established guidelines proposed by Kitchenham and Charters [82] and Garousi et al. [83]. The research protocol was modified and updated by conducting a pilot study of randomly selected 10 studies. The first author collected 90% of the data (62 out of 69), while the third author extracted the remaining 10% (7 out of 69). All data were shared in a collaborative folder and cross-checked by each author and any issues or disagreements were resolved in weekly research meetings among the authors.

While employing thematic analysis permits the data analysis to be grounded in the textual data collected from academic and grey literature studies, there is a threat of *subjectivity* of the data analysis [135]. To alleviate this threat, we discussed the issues and concerns of the emergent findings throughout the study in the weekly meetings. Throughout the iterative and intertwined rounds of data collection and analysis, the first author led the data analysis with support from other researchers who acted as validators at each stage.

3.6 Chapter Summary

From Chapter 2, we observed that current anti-phishing interventions encounter several obstacles, such as poor user interface design, lack of engaging and interesting content, incomplete or outdated anti-phishing instructions, flawed anti-phishing training implementation, and deficient anti-phishing policies within organizations. The usability issues that arise from the current one-size-fits-all anti-phishing interventions can be attributed to a need for greater awareness among practitioners of end-user requirements and preferences. There is a current lack of available personalized guidelines to assist practitioners for this purpose.

To address this gap, in this chapter, we first identified 22 dominant factors, consisting of 15 individual, 4 technical, and 3 organizational factors that impact the effectiveness and outcomes of anti-phishing interventions by analyzing the challenges and critical success factors reported in Chapter 2. This has enabled us to collect the factors required to tailor the design, implementation, and evaluation of the phishing interventions. We then devise 41 guidelines to aid practitioners in addressing the identified 22 dominant factors reported in this chapter and the identified 19 challenges reported in Chapter 2 within current anti-phishing intervention design, implementation, and evaluation. Our

guidelines are for four distinct practitioner groups: designers/developers, information security teams, cyber security experts, and C-suite employees. We offered guidelines for 14 different types of interventions within phishing education, training, and awareness interventions. Our personalized guidelines aim to improve the effectiveness of current anti-phishing software development, deployment, and assessment practices. By reporting these guidelines to address the needs of anti-phishing practitioners, we aim to contribute to the ongoing efforts to mitigate the threat of phishing attacks.

Statement of Authorship

Title of Paper	Understanding Practitioners' Challenges and Requirements in the Design, Implementation and Evaluation of Anti-phishing Interventions
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in a manuscript style
Publication Details	O. Sarker, A. Jayatilaka, S. Haggag, and C. Liu, M. Ali Babar, "Understanding Practitioners' Challenges and Requirements in the Design, Implementation and Evaluation of Anti-phishing Interventions," The Journal of Systems and Software.

Principal Author

Name of Principal Author (Candidate)	Orvila Sarker		
Contribution to the Paper	Performed analysis on all data, interpreted data, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	85%		
Signature		Date	20/12/2023

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Asangi Jayatilaka		
Contribution to the Paper	Supervised development of work, helped in developing methodology, data interpretation and edit the manuscript.		
Signature		Date	20/12/2023

Name of Co-Author	Sherif Haggag		
Contribution to the Paper	Supervised development of work, helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Please cut and paste additional co-author panels here as required.

Name of Co-Author	Chelsea Liu		
Contribution to the Paper	Supervised development of work, helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Name of Co-Author	M. Ali Babar		
Contribution to the Paper	Helped to evaluate and edit the manuscript.		
Signature		Date	20/12/2023

Chapter 4

Practitioners' Challenges in Practice and their Perspectives on the Personalized Guidelines

Related Publication: This chapter is based on our paper: *"Understanding Practitioner's Challenges and Requirements in the Design, Implementation, and Evaluation of Anti-phishing Interventions", submitted to The Journal of Systems and Software.* [CORE ranking: rank A].

In Chapter 3, we devised 41 guidelines from the literature (Chapter 2) to aid practitioners in addressing socio-technical challenges in the design, implementation, and evaluation of anti-phishing interventions. Unfortunately, the practical implementation of these guidelines to protect end-users and organizations may have its barriers. Until practitioners' challenges, requirements, and preferences in practice are better understood, it is not clear how the proposed guidelines can be effective in practice. To date, no study has investigated the challenges practitioners face in the design, implementation, and evaluation of anti-phishing interventions and no study has evaluated the practitioners' challenges in implementing the guidelines proposed in the literature. To address this gap, we conducted 18 semi-structured interviews with anti-phishing intervention and tool designers, security practitioners involved in the design, implementation, and evaluation of anti-phishing interventions, and C-suite employees involved in decision-making from 18 organizations in 6 countries. This chapter discusses the following contributions of this thesis: (1) we present a holistic overview of the current anti-phishing practices in the organizations; (2) we identify 8 challenges including challenges in phishing training content design, limitations in anti-phishing datasets, resource constraints in the organization for anti-phishing defense, and challenges associated with post-training phishing knowledge assessment of the employees in the organizations. We perform a comparative analysis of these challenges with the challenges we identified from the literature (discussed in Chapter 2) to demonstrate the ecological validity of the challenges identified from the literature. Based on the analysis we provide a set of recommendations to overcome these challenges; (3) we report practitioners' feedback and insights on our guidelines (discussed in Chapter 3) to understand their usefulness and practicability. We generate several findings that can provide insights to future researchers on how to make their devised guidelines applicable

and useful for the practitioners; (4) we gather practitioners' preferences and recommendations regarding an envisioned tool to access these academic guidelines. Based on this study, we provide four recommendations to bridge the gap between industry and academia in anti-phishing research and to improve the anti-phishing practices in organizations.

4.1 Introduction

Email and anti-phishing measures are regarded as a socio-technical system wherein security experts, end users, and technologies influence one another, yielding synergistic consequences [136]. To design effective phishing countermeasures that effectively align with users' requirements, it is crucial to understand anti-phishing endeavors as a collaborative process involving security experts, end users, and the technologies in subject [136], [137]. As phishing targets individual users, research has suggested anti-phishing solutions to be user-centered on top of using automated email filters, regular security updates in software systems, scanning malware presence, etc. [12]. Phishing education, training, and awareness interventions are considered a strong defense mechanism against phishing [14], [15].

The effectiveness of phishing education, training, and awareness interventions largely depends on the decisions and choices made by the anti-phishing intervention developers and practitioners [137], [138]. As observed in Chapter 2, the existing body of research reported several usability issues of phishing interventions that stem from the failure of the practitioners to duly account for end-users' requirements and preferences in the design, implementation, and evaluation phases. For example, anti-phishing toolbar designers are missing a significant amount of practical and useful information required for the user for the toolbars [42]; email client providers do not consider deploying critical security indicators required for unverified incoming emails [43]; browser warning designers do not follow the recommended guidelines for mobile browser phishing warnings which is putting end-users at phishing risks [52]; browser phishing warning design failures result in a low warning adherence rate as users do not properly understand the warnings [53]; poorly designed phishing training content by the practitioners leading to ineffective training programs in changing employees' behavior [54] and so on. These examples indicate that practitioners need to be aware of users' requirements and preferences to improve the usability issues of these interventions. Unfortunately, there is a dearth of guidelines available to assist practitioners in navigating the socio-technical challenges associated with integrating a spectrum of end-user requirements. Consequently, users of these interventions remain vulnerable to phishing attacks.

To support anti-phishing developers and practitioners, some recent studies provide brief recommendations. Brunken et al. [66] provided 7 recommendations to minimize the cost associated with conducting and running phishing simulation campaigns based on a case study in a single organization. To establish effective organizational awareness, Hillman et al. [63] provided recommendations for the information security officers of the organizations based on a case study in a financial institution in Israel. Based on a large-scale long-time

experiment in an organization, Lain et al. [22] suggested organizations conduct a prior investigation before adopting phishing prevention tools. None of the aforementioned recommendations considered guiding practitioners in addressing socio-technical issues and incorporating user requirements and preferences in the design, implementation, and evaluation of phishing interventions. To address this gap, we have devised 41 guidelines (discussed in Chapter 3 [139]) for 4 practitioner groups to support the design, implementation, and evaluation of 14 different interventions produced from critical success factors identified from the existing academic and grey literature (discussed in Chapter 2 [71]).

Although academic guidelines offer insightful information, they are rarely used by practitioners in the industry [62]. Research has shown that academic findings are abstract, complex, or too uncertain, which limits the ability of practitioners to utilize these findings in practice [62], [140]–[142]. Often academic recommendations are not presented in a format that meets the preferences of the practitioners and do not match the design practices used in the industry [142], [143]. Understanding practitioners' perspectives on these guidelines can facilitate translating and incorporating their perspectives into the guidelines [60]. This motivates us to understand the practicability and usefulness of our guidelines to address the socio-technical issues in practice and to investigate the factors that need to be considered to improve these guidelines' adherence in practice. Therefore one of our objectives in this chapter is to investigate practitioners' perspectives on our guidelines.

Again, very few studies highlighted the challenges and implementation difficulties practitioners face regarding the design, implementation, and evaluation of phishing interventions in real-world practice. For instance, Raffetseder et al. [65] documented the technical challenges associated with the porting of an anti-phishing plug-in designed for Firefox to Microsoft Internet Explorer. Despite the conceptual similarities between these two browser plugins, the process of adaptation posed significant technical impediments; a study conducted by Althobaiti et al. [47] has revealed that sudden phishing campaigns pose a challenge to the help desk employees of the organization causing delays in handling phishing reports. A recent study has discussed the procurement and adoption challenges involved with phishing simulation campaigns including extra time and effort required from the practitioners involved [66]. Therefore, the absence of available guidelines discussed in the previous paragraphs constitutes merely one facet of the obstacles requiring attention. A comprehensive understanding of practitioners' challenges and requirements in navigating the intricacies of designing, implementing, and evaluating the interventions necessitates investigation as well. To date, there is a lack of investigation specifically addressing the challenges encountered by practitioners involved in the design, implementation, and evaluation of different phishing education, and training awareness interventions. Therefore one of the objectives of this chapter is to identify the challenges faced by practitioners in designing, implementing, and evaluating phishing interventions.

Our 20 challenges in phishing education, training, and awareness interventions identified through the systematic multi-vocal literature review (discussed

in Chapter 2) are distributed among various interventions. For example, Challenge 1 (*UI design restrictions in the browser and email client*) and Challenge 3 (*Design constraints for anti-phishing warning UI interfaces*) are specific to phishing awareness interventions whereas Challenge 2 (*Content restrictions for phishing education and training*) is specific to phishing education and training interventions. Also, these challenges are spread across different stages of interventions. For example, Challenge 9 (*Anti-phishing technology deployment challenge*) is specific to the implementation stages and Challenge 20 (*Lack of post-training user knowledge retention practice*) is specific to the evaluation stages. Again, the human-centric, technical, and organizational factors we have identified (e.g., *age*, *complacency*, *confusion*, *curiosity*, *gamer type*, *organization position*), as discussed in Chapter 3, are specific to certain challenges in certain interventions. For example, the factor *gamer type* is relevant to phishing educational game-based interventions, and the factor *distraction* is specific to browser warning interventions. Inspired by the aforementioned illustration, we recognize a necessity for practitioners to access the guidelines linked with each intervention type, intervention stages, and socio-technical factors (presented in Chapter 3) in an organized and personalized manner. This consideration leads us to an additional objective: the identification of features preferred by practitioners in a tool designed to deliver guidelines in an organized and personalized manner. The features extracted through our investigation can be incorporated into a tool aiming to streamline practitioners' search for guidelines tailored to their specific requirements, thereby optimizing time efficiency.

4.1.1 Research questions

The main objectives of this chapter are to (i) obtain an overview of the anti-phishing defense mechanisms employed in practice (ii) to identify the challenges faced by the practitioners in designing, implementing, and evaluating phishing education, training, and awareness interventions (iii) evaluate our devised guidelines for the design, implementation, and evaluation of phishing interventions (discussed in Chapter 3 [139]) (iv) investigate the features required for a tool that present the guidelines in an organized and personalized manner.

The main research question we are seeking to address in this chapter is "**RQ5. What are the challenges, requirements, and preferences of the practitioners in adopting the identified guidelines in practice?**". While the motivations for this research have been already discussed in the previous section, this section serves to articulate supplementary points that substantiate our investigation. We have formulated four sub-research questions to address the issues discussed in the previous section.

💬 RQ5.1. What are the current anti-phishing practices employed by organizations?

Motivation: Organizational structure plays an important role in facilitating user-centric design, implementation, and evaluation approaches [144]. Although security practitioners' decisions can influence employee behavior in an organization, employees nevertheless think and act according to the organization's structure, policy, and culture [145]. As a consequence, it is common

that security practitioners integrate users' perspectives into the design, implementation, and evaluation of security measures in an unsystematic manner [145]. Therefore, instead of blaming security practitioners, it is recommended to understand and conceptualize the organizational anti-phishing practices and structures to improve organizations' security measures and policy designs [136]. We first get an overview of anti-phishing practices across organizations through RQ5.1. An overview of anti-phishing practices can aid decision-makers in evaluating the necessity of formulating, enhancing, or enforcing prevailing cyber security governance frameworks, regulatory protocols, and standards on phishing prevention.

💬 ***RQ5.2. What are the socio-technical challenges perceived by the practitioners in designing, implementing, or evaluating anti-phishing interventions?***

Motivation: Research conducted in real-world settings highlights the intricate challenges associated with implementing anti-phishing interventions such as deploying phishing education and training programs [63], [64]. One of the main reasons behind this is the unique requirements of each organization; for example, small and medium enterprises might have different requirements due to resource limitations, fewer security officers and specialists to review the quality of anti-phishing defense mechanisms in their organizations [66]. Also, the contradictory findings between prior literature and common industry practices in terms of phishing [22] highlight the need to investigate the challenges in industry practices. This leads to our second research question. The main motivation of RQ5.2 is to identify practitioners' challenges in real-world settings in designing, implementing, or evaluating anti-phishing interventions. This would open opportunities for future researchers to conduct in-depth investigations of these challenges, consequently advancing toward the development of innovative solutions to effectively tackle the identified issues.

💬 ***RQ5.3. What are practitioners' perspectives on our guidelines for designing, implementing, or evaluating anti-phishing interventions?***

Motivation: It is recommended that the guidelines, models, and frameworks suggested in the literature should be evaluated to understand and investigate their usefulness and applicability in practice [59]. Recommendations without empirical evaluation, lack implementation details practitioners require to feel confident to adopt them [60], [61]. Not all academic research is suitable for dissemination to practice [60], [62]. Questions were raised if the results of the academic findings are transferable to an industry setting [48]. Large-scale studies on anti-phishing simulation and training in real-world settings have led to new or contradictory findings than previous studies conducted in lab settings or limited demographics [22], [63]. This demonstrates the importance of understanding this domain of phishing in large-scale settings. Our guidelines were devised from the academic and grey studies. Most of these studies conducted on phishing interventions lack clear external validity due to the small sample sizes [67]–[69], little diversity [68], [69], or role-playing scenarios [19], [68]. Therefore, RQ5.3 would help us to understand their applicability in large-scale settings.

💬 ***RQ5.4. What features are desired by the practitioners for a tool that intends to support them in accessing the available guidelines devised from the literature?***

Motivation: An important aim of this study is to present the available guidelines in an organized and personalized manner to practitioners. This can help also remove or minimize potential barriers that prevent or limit practitioners' ability to adopt and follow findings from scientific research [146]. One of the main barriers practitioners face to the use of academic research is accessibility issues of relevant academic resources [60]. Practitioners often do not know the correct search term to find relevant research findings [60]. Furthermore, time constraints faced by practitioners significantly limit their ability to conduct extensive searches and read through the academic literature. Paywalls also pose a challenge for the practitioners to access academic research [62]. We formulate RQ5.4 to understand how and in what format practitioners would prefer to access the anti-phishing recommendations.

4.1.2 Summary of our findings

In this chapter, we discuss the following contributions:

• We present an overview of current anti-phishing practices in organizations including the phishing education and training methods currently in place in the organization, the approaches taken to design the education training contents, and manual and automated mechanisms applied to protect the organization from phishing. This knowledge would assist responsible authorities and decision-makers in investigating and understanding the necessity to develop or refine current cybersecurity governance frameworks and regulations or mandate certain security practices and standards to ensure higher phishing protection across organizations.

• We report a list of 8 challenges practitioners face in the design, implementation, and evaluation of anti-phishing interventions. These challenges include obstacles in training content design, limitations of anti-phishing datasets, limitations of training materials, challenges to motivate employees to encourage secure behavior, etc.

We also compare and contrast the challenges identified in the literature (Chapter 2) with the challenges identified in this chapter to understand the alignment of challenges in practice with the challenges discussed in the literature. We observe several similar challenges both in literature and practice: the use of culturally biased phishing training content, challenges in identifying phishing emails originating from legitimate domains, difficulty in handling phishing reports due to the huge volume, the presence of bots affecting the phishing assessment results, etc. Our analysis reveals some new findings from practice that have not been covered in the literature: budget constraints for adopting phishing-as-a-service training tools, lack of expert internal staff required to hire external staff for policy implementation, extra cost, and high maintenance requirements for customizing training content, etc. Our analysis produces a set of insights (findings) for similar challenges on how to overcome these challenges.

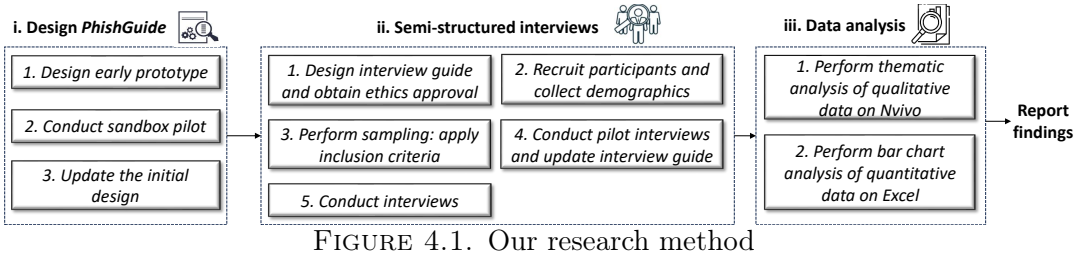

FIGURE 4.1. Our research method

• We gather perspectives from practitioners on our guidelines, aiming to discern their applicability and identify potential impediments to their effective implementation within real-world organizational settings. Our observations indicate that the guidelines are appreciated for their role-based categorization, inclusion of rationales, and utilization of high-quality sources during their formulation.

Our findings underscore the necessity of comprehending an organization's existing security measures, financial constraints, and established policies and procedures when implementing these guidelines. Several insights emerge from our study that can inform researchers on presenting and developing academic findings in a manner useful to practitioners. These insights include the incorporation of illustrative examples with the guidelines to enhance their utility for practitioners; the inclusion of relevant tools and technologies necessary for guideline implementation; the integration of statistical and demographic information employed in the studies informing the guidelines to improve reliability; and the importance of regular updates to the guidelines to ensure relevance in addressing evolving phishing threats.

• We extract desired features for an envisioned tool to access the guidelines by presenting a tool prototype *PhishGuide*. We collect a diverse set of features deemed desirable by the practitioners for a tool designed to access guidelines. These include options for both customized and non-customized guideline generation, a streamlined set of choices for guideline production, an input query-based system leveraging Natural Language Processing (NLP) to generate guidelines, automatic updating mechanisms for new content, and the provision of a comprehensive tool functionality summary for users. Our findings present an opportunity for future researchers and tool developers to incorporate practitioners' recommendations into tools aimed at assisting anti-phishing software developers and practitioners in accessing guidelines and recommendations.

4.2 Methodology

This section describes the different stages of our methodology. The study was approved by the Human Research Ethics Committee of the University of Adelaide. The ethics approval form can be found in Appendix F. Figure 4.1 provides an overview of our research methodology.

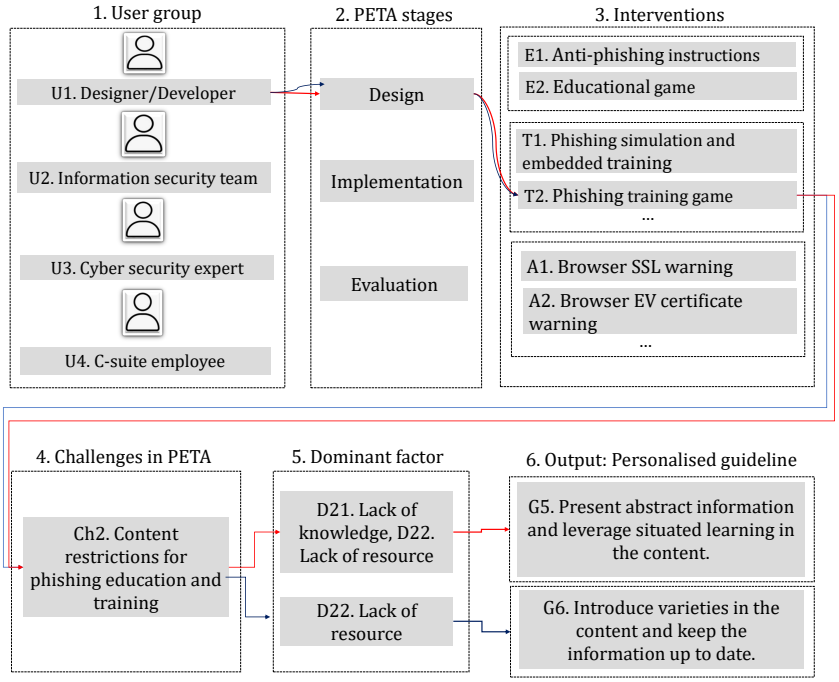

FIGURE 4.2. Example demonstration of early prototype

4.2.1 Design of PhishGuide

We designed a guideline generator tool *PhishGuide* (1) to present the guidelines to the participants in an organized manner and (2) to systematically collect participants' desired features for an intended tool that aims to generate personalized guidelines for them.

Early prototype

To design *PhishGuide*, we utilize the guidelines and the series of mapping described in Chapter 3. In Chapter 3, performed a systematic mapping among the: (i) 4 practitioner groups (designer/developer, information security team member, cybersecurity experts, and C-suite employees), (iii) 3 types of intervention stages (design, implementation, or evaluation), (iii) 14 types of interventions, (iv) 19 challenges in anti-phishing interventions, and (v) 22 human-centric, organizational, and technical factors (dominant factors). Our preliminary prototype was designed on Google form to collect these five pieces of information from the practitioners through 5 questions during their interaction with *PhishGuide*. Based on the answers provided by the practitioner, a maximum of two personalized guidelines were generated to assist them in their anti-phishing endeavors. The earlier prototype is available here [147].

Figure 4.2 depicts the earlier version of the prototype tool *PhishGuide*, which generates two guidelines (i.e., G5 and G6) based on two different sequences of inputs displayed and selected by the user (in this case, designer/developer). For the sake of brevity, the rationales along with the guidelines have been omitted from Figure 4.2 in step 6. Assume that a designer/developer (U1) wishes to design a phishing training game (T2). As identified in Chapter 2 and mapped

in Chapter 3, the challenge relevant to education games for the designers/developers is "content restrictions for phishing education and training (Ch2)". The corresponding dominant factors involved with these challenges are *lack of knowledge (D21)* and *lack of resource (D22)*. In *PhishGuide* interface, when the designer/developer enters the following sequences of inputs *U1. Designer/Developer → Design → T2. Phishing training game → Ch2. Content restrictions for phishing education and training → D21. Lack of knowledge, D22. Lack of resource*, *PhishGuide* generates the relevant guideline G5. Similarly, for the following sequences of inputs *U1. Designer/Developer → Design → T2. Phishing training game → Ch2. Content restrictions for phishing education and training → D22. Lack of resource*, the *PhishGuide* produces guideline G6. Hence *PhishGuide* displays personalized guidelines based the the sequences of options selected by the user. It is important to highlight that one single guideline (G5 in our example) can be capable of solving multiple human-centric issues (D21 and D22 in our example). This is grounded by the recommendations derived from the literature.

Sandbox pilot

Sandbox pilots are typically performed where researchers act as the participants in testing a tool that may still have issues regarding performance or usability. Sandbox pilots reduce the trouble of recruiting outsiders and help resolve issues before conducting a pilot with actual participants [148]. Three experienced software engineering researchers (who are not authors of the study) reviewed our early prototype. As discussed in Section 4.2.1, the initial design of *PhishGuide* only presents personalized guidelines, not all the guidelines. One improvement suggested by the researchers during the sandbox pilot was to allow the user of *PhishGuide* to access all the guidelines at a glance in addition to the personalized guidelines.

"*Just clicking "Next" without selecting any options apparently results in a complete breadth-first traversing of all nodes in the decision tree. So a user who just wants to see the list of all guidelines (not filtered by any or only some of the selection criteria) has to click "Next" 149 times to see the result: Guideline G1*" [Researcher 2].

Updating the initial design

These are the changes made on top of the initial design based on the findings of the sandbox pilot: (1) users of the tool are provided options to go through (i) all the guidelines, (ii) guidelines specific to their role, (iii) personalized guidelines automatically generated based on the answers provided on the 5 questions mentioned in Section 4.2.1, (2) we have also included a tool documentation option with a Google drive link for the participant to read the information used to design the tool which includes: methodological details of Chapter 3 [139] which was the basis of the tool, definition of the interventions etc. The updated tool can be found here [149].

TABLE 4.1. Demographic information of our participants

ID	Role	Location	Exp. (Year)	Degree	Organization domain	# of employees
U1	System designer	Australia	1 - 3	PhD	Scientific and industrial research	1000^{+}
U2	Member of security team	Australia	10^{+}	PhD	IT infrastructure	1000^{+}
U3	Member of security team	Australia	5 - 10	PhD	Software system and cybersecurity research center	50 - 249
U4	Member of security team	New Zealand	5 - 10	PhD	Farming	1000^{+}
U5	Member of security team	Australia	< 1	PhD	Energy	1000^{+}
U6	Software developer	Sri Lanka	1 - 3	Bachelor's	Transportation	10 - 49
U7	Security manager	Bangladesh	5 - 10	Master's	Legal	50 - 249
U8	UI/UX designer	Saudi Arabia	5 - 10	Master's	Education	50 - 249
U9	CEO	Australia	10^{+}	Master's	Risk analytics	10 - 49
U10	Member of security team	Australia	1 - 3	Master's	Cybersecurity strategy and innovation	1000^{+}
U11	CISO	Australia	10^{+}	Master's	Cybersecurity solutions and services	1000^{+}
U12	UI/UX designer	Australia	1 - 3	Master's	Global growth and operations	50 - 249
U13	Software developer	Australia	1 - 3	Master's	Education	50 - 249
U14	System designer	Indonesia	1 - 3	PhD	Education	1000^{+}
U15	Chief architect and CISO	Australia	10^{+}	Master's	Education	1000^{+}
U16	Software developer	Australia	< 1	Master's	Education	1000^{+}
U17	System designer	Australia	5 - 10	PhD	Education	1000^{+}
U18	Member of security team	Australia	10^{+}	Master's	Financial institution	1000^{+}

4.2.2 Semi-structured interviews

Interview protocol

The interview guide was designed to collect both quantitative and qualitative data. To understand the strengths and weaknesses of the guidelines, we designed several statements as close-ended questions by leveraging the existing study [150]. We adopted a Likert scale-based evaluation [151] for the close-ended questions. To evaluate usability (learnability, memorability, accessibility satisfaction, and efficiency [152]) and usefulness of the tool *PhishGuide* we leveraged the existing standardized questionnaire SUMI [153], a tool to evaluate the usability of a system. SUMI is built upon the definitions provided by ISO 9241-210 [154], the standard for requirements and recommendations for Human-centred design for interactive systems [155]. We also asked open-ended questions on the guidelines and tool as open-ended questions help identify the important concerns of the participants [156]. The interview questions can be found in Appendix E.

Recruitment

To recruit participants, we circulated a flyer using social media. The flyer contains information regarding the objectives of the study, the type of practitioners we are looking for, the task summary, a Google form link (or a QR code) to register for participation, the ethics approval number provided by our organization's human research ethics committee and our contact email. We also invited participants via email and social media messages.

Collection of demographic information

A Google form was designed to collect participants' consent and demographics, which includes their names, contact email, their roles in the organization, total industry experience, location of their organization, the organization domain, organization size and their highest level of education. The demographic form can be found in Appendix D. Table 4.1 shows the demographic details of our participants. The symbol "+" is used in Table 4.1 to indicate "more than", for example, 100^{+} refers to more than 100. Similarly, the symbol "<" is used to indicate "less than".

Participants are from different geographical locations: Australia, Sri Lanka, New Zealand, Bangladesh, Saudi Arabia, and Indonesia. Under the designer/developer group, we have anti-phishing system designer, software developer, and UI/UX designer. We have one C-suite employee: the CEO of an organization. Participant organizations operate in several different domains ranging from cyber security solutions and services, cyber security strategy and innovation, risk analytics, legal, transportation, energy, farming, IT infrastructure, scientific and industrial research, software systems and cyber security research, and education. Participants have experience ranging from less than 1 year (denoted as <1 in Table 4.1) to more than 10 years (denoted as 10^{+} in Table 4.1). The highest level of education reported by the participants are Bachelor's, Master's, and PhD.

Sampling

A total of 25 practitioners registered to participate in our study, and 21 of them met our inclusion criteria. We included participants who met one or more of the following inclusion criteria: (I1) Experience in designing, implementing, or evaluating phishing education, training, or awareness interventions or anti-phishing technologies; (I2) Head of the organization who are involve in the decision making. The inclusion criterion I2 was chosen to include C-suite employees, for example, the CEO of an organization. This is because these practitioner groups are involved in the policy-making, which can potentially affect the security of the organizations. Table 4.2 summarizes the experience of each participant. Please note that Table 4.2 only shows 18 participants who are the final set of participants after the pilot interviews discussed in the next section.

TABLE 4.2. Practitioner selection based on our inclusion criteria

ID	Practitioners' experience (described by the practitioners)	Criteria
U1	• I have designed anti-phishing technologies or solutions	I1
U2	• I have been involved in the incident response process for phishing attacks	I1
U3	• I have been involved in the incident response process for phishing attacks	I1
U4	• I participate in creating or implementing phishing awareness programs within my organization	I1
U5	• I participate in creating or implementing phishing awareness programs within my organization	I1
U6	• I have conducted user education sessions to improve phishing awareness	I1
U7	• I participate in creating or implementing phishing awareness programs within my organization	I1
U8	• I have designed anti-phishing technologies or solutions	I1
U9	• I have no experience dealing with or mitigating phishing attacks	I2
U10	• I participate in creating or implementing phishing awareness programs within my organization • I have been involved in the incident response process for phishing attacks	I1
U11	• I participate in creating or implementing phishing awareness programs within my organization • I have been involved in the incident response process for phishing attacks	I1
U12	• I have designed phishing awareness intervention	I1
U13	• I have designed phishing awareness intervention	I1
U14	• I have designed anti-phishing technologies or solutions	I1
U15	• I have experience in overseeing and managing cybersecurity teams and strategies within an organization • I participate in creating or implementing phishing awareness programs within my organization	I1
U16	• I have designed phishing awareness intervention	I1
U17	• I have conducted user education sessions to improve phishing awareness	I1
U18	• I have experience in overseeing and managing cyber security teams and strategies within an organization • I participate in creating or implementing phishing awareness programs within my organization	I1

Pilot interviews

Adhering to the recommended practices for empirical studies with human subjects [148], we conducted pilot interviews with three randomly selected developers to identify potential issues related to the study design. The main objectives of the pilot interview were (1) to enhance the clarity of the interview protocol and (2) to check whether a 40-minute interview was of sufficient duration to capture the details. Based on the pilot interview, we identified two issues: (1) initially participants had the (wrong) impression that they needed to evaluate all 41 of the guidelines, and (2) some participants misunderstood the questions: the challenges they were asked to describe were interpreted as the challenges faced by the end-users, rather than those faced by themselves. We updated the interview protocol based on the feedback received from the participants to avoid confusion. We discarded pilot data to avoid extraneous variations as suggested by the guideline [148].

Interviews

We conducted 40-minute semi-structured interviews on Zoom as the participants were from different geographical locations. Following an existing study [145], during each interview, we provided a brief overview of the guidelines and how the guidelines are devised based on the methodology discussed in Chapter 3. Then, we provided a live walk-through of the PhishGuide using the Zoom screen-sharing option. During the demonstration of the tool, we asked participants to select their preferred options while interacting with the tool. It is noteworthy to mention that, due to time limitations, participants were presented with a maximum of randomly selected 5 guidelines generated related to their roles through different combinations of options embedded in the tool. However, some curious participants voluntarily spent more than 1 hour discussing all of the

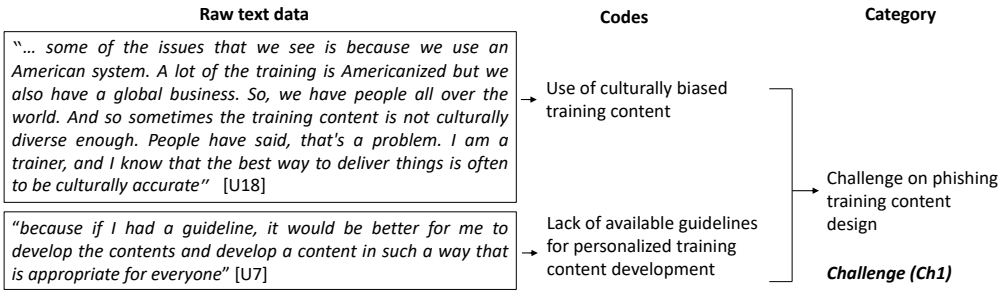

FIGURE 4.3. Our data analysis process to identify the challenges faced by the practitioners

guidelines relevant to their roles. Then, we started the interview by asking them open-ended questions on the current challenges and practices in the design, implementation, and evaluation of anti-phishing systems and interventions. We then shared a 15-minute Google form link consisting of close-ended questions in the Zoom chat box and requested participants to share their screens while completing the form during the interview. This allowed us to ask follow-up questions from the participants for each question they responded to. We asked follow-up questions when participants completed the form to obtain high-quality and rich data compared to the case if the participants were to send a survey form and complete it later [157]. All interviews were recorded and transcribed. To improve the accuracy, the transcriptions auto-generated by Zoom were manually checked and corrected by listening to the recorded audio.

4.2.3 Data analysis

For the qualitative data, we performed an *iterative*, *open*, and *axial* coding approach [106] using the qualitative data analysis software NVivo. Figure 4.3 shows an example of our thematic analysis process. Our interview transcripts were uploaded to NVivo and all of our research questions were coded with the same codebook. We continued coding until no new themes were identified. The first author analyzed the data and all codes were cross-validated by other authors to reduce subjectivity bias and improve the reliability of our findings as suggested by the best practices [158]. Any disagreements were discussed and resolved in the weekly research meetings. Initially, following existing research [12], the first author conducted a pilot data analysis with two interview transcripts, in order to understand the pattern of the data and to identify the emergent themes. The findings were discussed with other collaborators and modified based on the feedback. For the quantitative data, we employed a bar chart analysis, which is a recommended method of analysis for individual statements in the questions [159].

Empirical studies employing qualitative research methodologies commonly evaluate theoretical saturation to ascertain the adequacy of sample sizes in drawing meaningful findings [160]. We adopted a similar approach; however, we intended to understand whether the continuous addition of participants contributed to the generation of new themes. To demonstrate this, we utilized the data gathered for RQ5.2 (i.e., participants' perceived challenges) as a means

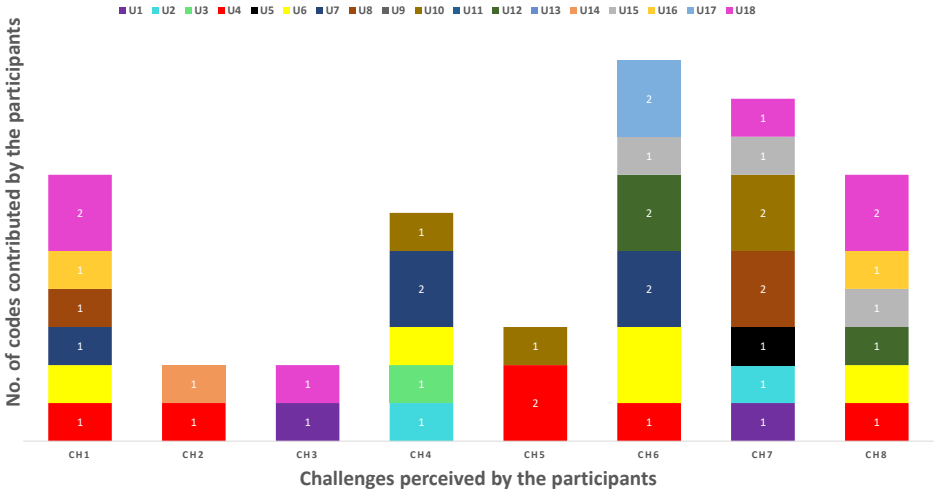

FIGURE 4.4. Assessing saturation for sample size selection

to gauge the point at which theoretical saturation could be achieved. Figure 4.4 indicates that beyond participant U14, no new themes emerged. After U14, each instance of data analysis involving a new participant adds a code to the preexisting set of identified challenges.

4.3 Research Findings

Findings of our four research questions are presented in Section 4.3.1, Section 4.3.2, Section 4.4.1 and Section 4.4.2. Symbols U[*] and P[*] used in this chapter represent the practitioner ID (please refer to Table 4.1) and the primary study ID (these are the studies collected for our MLR discussed in Chapter 2) respectively.

4.3.1 Anti-phishing practices in the industry

This section presents the findings of RQ5.1.

Phishing education and training methods

We observe that organizations conduct different forms of phishing education and training to raise awareness among employees. In large organizations, phishing simulation is a common technique where employees are sent random simulated phishing emails during working hours to test their employees' phishing susceptibility. The employees who do not perform well in the simulation test receive video-based training. Mostly, these simulations are run by the organization itself or at a government level. Only one of our studied organizations considers any embedded training soon after the employees click on these simulation emails. In small organizations, a common method used is induction-level training, where new employees receive basic instructions on phishing. Some organizations do not have any mandatory training in place for their employees, staffs volunteer

their time and effort to educate other employees: "*In my organization, at least it is more like voluntary work like lecturers or professors who are teaching the cyber security courses are the people who are giving training to...the employees*" [U8], "*we ran campaigns in [university name] and different other colleges to increase the awareness of social media users and the challenges that come through social media via phishing links*"[U7].

Phishing education and training contents

We notice that large organizations mainly use video-based interactive training materials to train and educate their employees. These training materials introduce employees to the common anti-phishing terminologies, certain words, and clues to look for and what needs to be done to protect themselves from phishing: "*in terms of content I would say, it's pretty interactive... for example, concrete scenarios in daily lives of what certain you know terminologies mean and what they look like*" [U14]. Training contents are personalized based on employees' roles, types of attack vectors, the domain of the organizations, and the associated risks in the organization: "*we have a couple of different templates we design based on the kind of attacks we see, and we heavily target the high-end people who are most likely to get those phishing emails, so everybody gets it. But we watch over those ... high-risk users. So we take care of them much more carefully*" [U4]. Some security practitioners do their own research, use their own experience, and take suggestions from experts to add real-world phishing examples to make the training content interesting and engaging: "*I develop the content by myself. But I talk with, say, with others ... like I have connections with [organization name]... so I actually share my slides with them and they also share some information*" [U7].

Phishing education and training frequency, reminders, and notifications

Our investigation reveals that organizations run phishing simulations both (1) randomly without sending any pre-notification, and (2) systematically by sending pre-notification a few weeks before running the phishing simulation: "*admin team... beforehand, like one month prior, they give notice to all the employees that there will be phishing simulation run on in one month's time. So that's the only prior notice. It teams to schedule this email to be sent out to all the employees*" [U5]. In terms of training frequency, some organizations run training sessions quarterly and analyze the results after each quarter. While others run phishing education sessions whenever there is an available slot in between other training programs conducted in the organizations: "*So they give us a slot almost in every training program, like, for example, if there is a training program going on money laundering ... we ask them to give us 1-hour or 2-hour slots so that we can share the concept of phishing and we can connect the money laundering issue with phishing*" [U7]. Employees are sometimes sent text-based anti-phishing instructions via email as a reminder of what to do to protect themselves from phishing: "*They will just give some information ...*

for example, these are the things that need to be done to avoid phishing, and then it will give 3 bullet points" [U12].

Feedback and follow-up on phishing education and training

It is evident from our data that none of the organizations included in our study follow-up with the employees on their performance on the phishing simulation. Moreover, these organizations also do not collect feedback from the employees on the education and training materials used by the organizations: "*no, we don't get feedback on the actual simulation and the actual training the only feedback that we get is if they click it, or if they don't click it, and if they pass, then we have, like an automated process to go outside. if they fail, then it then gives them that training*" [U10].

Employees' phishing knowledge assessment

After educating employees, some organizations do not test their knowledge level on phishing. This is common for the organizations that conduct instructor-based phishing educational programs or provide only inductions level training: "*No, till now we do not have any policy to test their efficiency, even we suggest sometimes that we should have a red team blue team campaign to check the efficiency*" [U7]. Organizations that run phishing simulation programs analyze employees' behavior and susceptibility after each simulation session, and based on the results, they invite employees for further training: "*we can track everything like who's opening who? What time they're opening? What sort of browser they're doing, you know, they're all set up*" [U4].

Manual and automated phishing detection and prevention mechanisms

Large organizations use automated anti-phishing solutions deployed in the email clients (e.g., Mimecast [161]). Some organizations take a manual approach where security team members manually analyze some metadata, such as the email headers, to identify inconsistencies, spoofed sender addresses, or unusual routing patterns to look for phishing attempts. Small organizations depend on cloud computing platforms and services (e.g., Microsoft Azure) to protect their organization against phishing: "*we depend on the security levels and measures that Azure would provide in order for us to get protected other than that, we don't have any sophisticated phishing mechanisms in place*" [U9]. Prior investigations are performed in the organizations before adopting any anti-phishing tool or interventions and a suitable one is chosen based on the budget: "*well, the first major thing that comes into it is budget which kind of sucks... So then we'll nail it down to well, let's just say 4 options. And then after that, we'll do an extensive analysis on those. And then we can do like an options matrix and then sort of work out which one's best fit for us*" [U10].

Actions taken on real phishing incidents

After the occurrence of an actual phishing attack, employees are alerted through an emergency meeting and they are reminded of the basic hygiene factors again: "*we had an immediate meeting stating that this happened. And you need to be careful with all that stuff... So they just give a call for each team member, and then they'll just give it like a meeting. and give like information*" [U12]. Some organizations send email notifications to the employees to let them be aware of the attack: "*And we actually send, take a screenshot and send. Okay, this is something that we got. It looks like this is not a good thing. If you have something like this, it is better to block*" [U6]. If any employees fall victim to actual phishing attacks, some organizations have help desks to support the employees. They have incident response plans and phishing playbooks, which is a step-by-step guide they follow after the attack: "*Yes, so we have a help desk. we do have an incident response plan. And we also have an incident response, phishing playbook*" [U10].

4.3.2 Challenges in the design, implementation, and evaluation of anti-phishing interventions

This section presents the findings of RQ5.2. Table 4.3 shows the challenges we derived from our interview data with the key points.

Challenge 1. Challenges on phishing training content design

One of the main challenges in training content design is the lack of available guidelines for practitioners to design personalized training content: "*because if I had a guideline it would be better for me to develop the contents in such a way that is appropriate for everyone*" [U7]. Organizations often use outdated training content, which is not really helpful in changing employees' security behaviors. For the anti-phishing intervention designers, finding a specific time frame to update the content is another challenge. It is hard to find an appropriate time interval for updating the content incorporated in the intervention, as phishing threats and tactics are constantly evolving. Anti-phishing interventions designed to suggest suitable resources to the users based on their phishing knowledge are facing challenges in recommending useful anti-phishing instructions to the users as sometimes these resources are not accessible to the users: "*So our initial idea was to suggest them [users] some resources. But then, like most of the resources, they're not free or fully accessible to the users. So we provided some free online resources for each topic to enhance their skills*" [U16].

Challenge 2. Lack of available phishing datasets

There is a scarcity of phishing website datasets required to train the anti-phishing tools: "*there's really no dataset of these websites. So I have to scrape that by myself, collect that daily, you know, make a script to ultimate*

that process" [U14]. Also, many of the available datasets do not cover data for the attack types, for example, SMS-based phishing attacks - smishing.

Challenge 3. Limitations of anti-phishing datasets

Phishing detectors are often trained on legitimate URLs, but this training methodology leads to a problem. Many phishing detectors end up misclassifying legitimate sites as false positives. Available anti-phishing datasets to train the anti-phishing tools have limitations arising from class dependency. Certain domains are targeted more frequently than others due to the prevalence of digital platforms and e-commerce. Not all e-commerce websites have efficient target URLs or web pages associated with them, making them vulnerable to attacks. "*most of the phishing detectors are trained in a way that you have legitimate URLs or sites. Then, there are some small subsets of those legitimate sites that have phishing URLs or web pages associated with them. Because of that, I found most of the phishing URL detectors or phishing detectors detect legitimate sites as false positive*" [U1].

Table 4.3. Challenges in anti-phishing interventions in practice

Challenges	Key points (included practitioners' ID)	#
Challenge 1. Challenge on phishing training content design	• Challenging to design training content to change user behavior [U4] • Difficult to select a time frame to update the contents of the intervention [U16] • Lack of available guidelines on personalised training content development [U7] • Lack of personalised training content [U8, U18] • Limited free online anti-phishing resources for the end users [U16] • Restricted use of bands for training content development [U18] • Use of culturally biased training content [U18]	5
Challenge 2. Lack of available phishing datasets	• Unavailability of dataset that covers all types of phishing attack [U14]	2
Challenge 3. Limitations of anti-phishing systems and interventions	• Issues in the reporting systems reduce the reporting rates [U18] • Lack of interactive and explainable interventions [U1, U4] • Anti-phishing tools detection mechanism is biased towards popular phishing trends and platforms [U1] • Phishing detectors trained with legitimate sites where a subset of phishing URLs gives false positives [U1]	2

Continued on next page

Challenges in anti-phishing intervention in practice – continued from previous page

Challenges	Key points (included practitioners' ID)	#
Challenge 4. Challenge due to complicacy of phishing attack	• Phishing emails originating from legitimate domains present additional challenges [U7] • Difficult to automate incident response due to the presence of unforeseen threats [U10] • Notifying employees during the incident automatically is not possible [U6] • Automated filters deployed in the email clients generate false positives [U7] • Conscious employees click on phishing links when they have heavy workload [U3] • Manual rule-based email filtering is challenging [U2]	5
Challenge 5. Challenge in striking a balance between training frequency and security fatigue	• Difficult striking a balance between training frequency and minimizing security fatigue [U4, U10]	2
Challenge 6. Lack of motivation and attention among employees	• Employees use weak email passwords [U7] • Employees visit unsafe sites using personal devices [U6] • Extra effort is required to educate new employees [U6] • Lack of attention by the employees [U4, U7, U12, U17, U15] • Employees lack willingness to attend phishing training sessions [U12, U17]	5
Challenge 7. Resource limitations of the organizations	• Difficulty in handling phishing reports due to huge volume [U1, U2] • Need additional time and effort for anti-phishing training and education [U5, U10, U15] • Budget constraints for adopting phishing-as-a-service training tools [U8, U10] • Lack of expert internal staff requires hiring external staff for policy implementation [U8] • Extra cost and high maintenance requirement for customizing training content [U18]	6
Challenge 8. Challenge on post-training phishing knowledge assessment	• Difficulty to determine if the anti-phishing instructions provided to the employees were useful [U6, U12] • Employees do not retain the anti-phishing knowledge taught to them [U4, U15] • Difficult to design questions to evaluate users phishing knowledge [U16] • Presence of bots affects the phishing assessment results [U18] • Challenging to avoid ceiling effect [U18]	6

Challenge 4. Challenge due to complicacy of phishing attack

Due to the complexity of the ways and tactics attackers use to initiate and deploy phishing attacks, it is often difficult to deal with these attacks. For example, phishing threats are constantly evolving, therefore it is hard to automate the incident response process: "*the incidents are always gonna be different every time. So it's kinda hard to automate something that you don't know*" [U10]. Sometimes attackers gain unauthorized access to websites or email

accounts, a situation that often goes unnoticed by employees or organizations immediately. Emails originating from these compromised legitimate email accounts pose a challenge for fellow employees in discerning their authenticity: "*in the past, I have seen some of our [department name] domain email addresses hacked. Then, those email addresses were used to share those phishing links. So when they actually compromise a legitimate email address that we trust and if we get an email from that domain with a phishing link, for general people, it's very hard to identify whether it's a phishing link or not because it came from a very legitimate email address*" [U7]. Often, employees click on phishing links in the emails that they receive during business hours due to a heavy workload. Again, employees' email accounts compromised during weekends or after business hours pose another challenge, as it is not always possible to instantly notify other employees about the attack outside of business hours. Practitioners also encounter challenges in safeguarding email clients through manual email filtering based on strict rules. Implementing stringent rules may result in overlooking crucial emails, while employing more lenient rules could increase security vulnerabilities: "*So if you use a very strict rule to remove those risks, it might lead to the fact that they [employees] missed some email or missed some information. However, if you relax a restriction, some security risks might get into our it systems*" [U2].

Challenge 5. Challenge in striking a balance between training frequency and security fatigue

Security professionals encounter challenges in finding an equilibrium between the frequency of training sessions and the reduction of security fatigue among employees. While it is crucial to train individuals who repeatedly violate security protocols, it is equally difficult to prevent them from becoming irritated due to excessive training sessions: "*which is a difficult thing for security people, is that you cannot interrupt the users, you know. You'll have to work in the background. So the minimum intervention that you'll need to have like if you make the users unhappy, you know, or they get annoyed*" [U4]. "*it's a bit of a tricky one because we don't want to make anyone feel bad for failing and we also don't wanna just make it a thing like, Oh, here's another security training that you have to do, and then, just like quickly go through it*" [U10].

Challenge 6. Lack of motivation and attention among employees

In section 4.3.1, we described that many organizations conduct voluntary training sessions with the aim of instructing their employees on phishing. The training materials provided, particularly in video format, are frequently skimmed through by employees, driven by the sole objective of completing them expeditiously. A pivotal factor contributing to this apathetic attitude and lack of motivation among employees arises from the absence of real-life examples within the training content, rendering it unrelatable and uninteresting for the workforce. This gives rise to security risks, including the use of organization email accounts for external services on unsafe sites and the adoption of weak passwords. "*so this person's email got hacked...And he has used the same password. They*

[attackers] had used this email to send some scams to all of his contacts. And he was locked out of his email, and they had to send 11 thousand or 15 thousand, that kind of amount... you can have all these software techniques. But we can't prevent them [employees] doing, you know, doing their personal devices. And it's not something that we monitor, but we had given them advice, saying, not to use a company email to register for outside services" [U6].

Challenge 7. Resource limitations of the organizations

Our participants have shared several resource constraints they face in their organizations: (i) handling phishing reports requires additional time and effort due to their substantial volume; (ii) training employees in batches takes a lot of time and coordination as in large organizations there are hundreds of employees; (iii) organization faces budget constraints for adopting anti-phishing tools and anti-phishing training programs; (iv) extra effort and attention requires to educate new employees; (v) lack of experts inside the organizations for policy implementation which requires them to hire external experts; (vi) due to lack of expert employees for a particular security task, practitioners often prioritize spending time and effort dealing with other security issues in the organization over phishing.

Challenge 8. Challenge on post-training knowledge assessment

Practitioners encounter several challenges concerning the assessment of employees' knowledge both during and after phishing education and training programs. A significant issue arises post-training, as employees often struggle to retain the information and instructions provided to them. Consequently, they fail to adhere to security policies and essential cyber safety practices: "*But what happens is that many users don't retain that. So after a few weeks, they might forget it, or they might be in a situation where they're not checking the details of the links or the domain of the email*" [U4]. Additionally, crafting appropriate questions to assess employees' knowledge levels presents a formidable challenge as phishing threats are constantly evolving. Furthermore, practitioners express a lack of effective methods to determine the efficacy of the instructions (e.g., email-based anti-phishing) imparted to employees: "*it's difficult to assess whether the instructions are clear to them because it's an email, right?*" [U6]. Moreover, numerous organizations lack formal protocols to ascertain whether employees have thoroughly comprehended the video training materials or if they have hastily completed them: "*if the person is completing it within 5 minutes, we know for sure that he just skipped everything if they completed for 15 minutes, we know they had read something*" [U12].

4.4 Comparative Analysis of the Challenges Identified in Literature and Practice

In this section, we analyze the challenges identified in the literature (discussed in Chapter 2) and the challenges perceived by the practitioners in this chapter.

This analysis would enable us to understand the ecological validity of the challenges we found in the literature in real-world contexts. Table 4.4 demonstrates a comparison of the key points in the challenges identified from literature shown in Table 2.5 and key points in the challenges shown in Table 4.3. From Table 4.4, we observe that most of the challenges identified from the literature are also discussed by the participants in this interview study. Table 4.4 also highlighted some new challenges identified in our interview study that were not discussed in the literature. We briefly discuss some insights from the mappings of the challenges shown in Table 4.4 based on the order they are included in the table:

TABLE 4.4. Mapping of the challenges identified from the literature and the challenges found in the organizations

No.	Findings from the Organizations	Findings from the Literature
	Similar Findings	
1	• Lack of available guidelines on personalised training content development [U7]	• Inadequate policies and guidelines to invoke user behavioral change [P38]
2	• Lack of personalised training content [U8] • Use of culturally biased training content [U18]	• Disregard for user misunderstandings and interests [P11,P19] • Disregard for both casual and serious gamers [P36] • Presence of cultural bias in the content [P36]
3	• Lack of interactive and explainable interventions [U1, U4]	• Lack of engaging and interesting phishing education and training material [P10, P19, P28] • Lack of comprehension and explainability [P14, P25, P49]
4	• Anti-phishing tool detection mechanism is biased towards popular phishing trends and platforms [U1] • Phishing detectors trained with legitimate sites where a subset of phishing URLs gives false positives [U1] • Automated filters deployed in the email clients generate false positives [U7]	• False positives and lack of reliability [P1, P2, P3, P8, P10, P13, P14, P18, P24, P25, P28, P44, P49, P57, P69]
5	• Phishing emails originating from legitimate domains present additional challenges [U7]	• Difficulty in detecting spear phishing due to personal relevance and familiarity [P1, P7, P14, P15, P21, P26, P49, P58]
6	• Notifying employees during the incident automatically is not possible [U6]	• Complicacy to safeguard employees in distributed and siloed settings due to enlarged attack surface [P6, P54, P57, P62, P65]
7	• Conscious employees click on phishing links when they have heavy workload [U3]	• Users' distraction by other tasks is not well considered [P2, P7, P8, P13, P14, P24, P47]
8	• Manual rule-based email filtering is challenging [U2]	• Training email spammed by email provider [P28]

Continued on next page

Mapping of the challenges – continued from previous page

No	Findings from the Organizations	Findings from the Literature
9	• Difficult striking a balance between training frequency and minimizing security fatigue [U4, U10]	• Frequent exposure causes warning fatigue [P4, P13, P14, P17, P18, P26]
10	• Employees use weak email passwords [U7] • Employees visit unsafe sites using personal devices [U6] • Lack of attention by the employees [U4, U7, U12, U17, U15]	• Ignorance due to lack of trust and understanding on phishing warning and training [P1, P2, P3, P4, P8, P11, P14, P24, P28, P31, P36, P39, P44, P49] • Disregard to warning due to appealing web content and site reputation [P2, P8, P14, P24, P49]
11	• Employees lack the willingness to attend phishing training sessions [U12, U17]	• Poor practice of training execution [P12, P59] • Requirement of users' effort and willingness to use anti-phishing warnings [P19, P31, P45]
12	• Difficulty in handling phishing reports due to huge volume [U1, U2] • Need additional time and effort for anti-phishing training and education [U5, U10, U15]	• Handling phishing incident reports requires the need for human validation [P50] • Embedded training deployment requires manual human effort [P45]
13	• Employees do not retain the anti-phishing knowledge taught to them [U4, U15]	• Effectiveness of phishing interventions subject to dwindle over time [P13, P21, P40, P45]
14	• Difficulty to determine if the anti-phishing instructions provided to the employees were useful [U6, U12] • Difficult to design questions to evaluate users phishing knowledge [U16]	• Lack of investigation on users' long term behavior change [P7, P31, P34, P54]
15	• Presence of bots affects the phishing assessment results [U18]	• Difficulty in measuring user phishing training effectiveness due to presence of bots [P55]
New findings from the organizations		
• Limited free online anti-phishing resources for the end users [U16] • Extra effort is required to educate new employees [U6] • Restricted use of bands for training content development [U18] • Issues in the reporting systems reduces the reporting rates [U18] • Challenging to design training content to change user behavior [U4] • Difficult to select a time frame to update the contents of the intervention [U16] • Unavailability of datasets that cover all types of phishing attack [U14] • Difficult to automate incident response due to the presence of unforeseen threats [U10] • Budget constraints for adopting phishing-as-a-service training tools [U8, U10] • Lack of expert internal staff requires to hire external staff for policy implementation [U8] • Extra cost and high maintenance requirement for customizing training content [U18] • Challenging to avoid ceiling effect [U18]		

It has been observed both from our MLR and the interview study that the current policies and guidelines in place within the organizational settings are not enough to positively influence employees' security behavior [P50, U7]. As highlighted in our primary study P38: "*while prior research has discussed the importance of security policies and standards in driving systematic compliance*

for the employees, we have noted that these policies and guidelines are not adequate to invoke behavioral change in employees dealing with phishing e-mails". This data has also been substantiated by our participant U7, who discussed the limited availability of guidelines regarding the design of content tailored for diverse user groups: "*because if I had a guideline it would be better for me to develop the contents and develop content in such a way that is appropriate for everyone...As there are no guidelines, I do not have any specific way of looking into that plan. So it's a very big challenge for me...*".

Insight: Organizations may collect feedback from employees to incorporate their experiences and insights into policy enhancements.

User expectations and requirements are not well incorporated in the design and execution of phishing education and training. Certain organizations employ culturally biased training content for phishing training, leading to numerous complaints from the employees. This tendency is particularly noticeable when an organization runs a global business with multiple offices worldwide, resulting in the appropriateness of the training content varying across different regions as mentioned by participant U18: "*some of the issues that we see is because we use an American system. A lot of the training is Americanized but we also have a global business. So we have people all over the world. And so sometimes the training is not culturally diverse enough. People have said, that's a problem. It's an issue in that. I am a trainer, and I know that the best way to deliver things is often to be culturally accurate*". Study P36 has discussed this cultural bias in training content: "*There were a number of issues and criticisms of APP [phishing game discussed in the study], one of which was the cultural bias within the game – namely that many of the websites in question were for American companies and thus unfamiliar to those outside of the United States. As a result, some participants found it difficult to assess whether aspects of the URL were suspicious or part of a company name (e.g. "CitiBank" where participants were unsure whether Citi was part of the name or a misspelling)*". Moreover, the literature also delves into other forms of bias, such as the utilization of game-based training tailored for either casual or serious gamers [P36].

Insight: Organizations are encouraged to employ training materials customized to the geographic location of the organization.

Certain limitations of current phishing interventions such as the lack of engaging and interesting training content are discussed in the literature which we also observed in our interview data. Studies [P10, P19, P28] have discussed that conventional phishing materials usually consist of some instructions or videos of what to do to cope with phishing attacks, for example: "*Anti-phishing education often struggles to capture the interest of end users. Materials commonly used for cyber security training include notes, videos, and email bulletins. However, these materials are often not very engaging and separate the learning material from the context in which employees routinely apply this information (e.g. email clients)*" [P19]. End-users do not find these materials interesting or engaging

enough and eventually do not learn the lessons taught to them. These findings are confirmed by our interview data. Participant U4 shared many big companies like theirs use non-interactive training materials and outdated training content which is not helping change the behavior of the employees: "*So it's I quite think, about like the tools we have. You know, we have very substantive tools, but the education stuff is still, I think, very outdated, like most of the big companies. It's normal; you either watch a video, then have some questions, and you answer the questions, and then you're done. But it doesn't help the users to change their behavior like, say, for example, if you read a manual and you feel like you can, you know, learn how to drive, you cannot*" [U4]. Although gamed-based training has been discussed as effective for improving learning outcomes and user satisfaction (e.g., P19), we observed that none of our studied organizations have adopted game-based training.

> **Insight**: Organizations may adopt game-based training for their employees to enhance overall outcomes.

Most of our studied organizations have automated anti-phishing tools deployed into their organizations. Most often these automated solutions provide the first line of defense for safeguarding employees from phishing attacks. Unfortunately, evidence from the literature and our interview study indicate that these automated anti-phishing tools can not be fully relied on due to the false positives generated by these solutions. The primary factor, as mentioned by participant U1 in our discussions, is rooted in the class-dependent nature of datasets employed for training automated phishing detection systems, coupled with the limited accessibility of target URLs or web pages affiliated with less frequented websites: "*I found many challenges... like into a class dependency, for example, there are certain domains that are phished more often than the other domains. And there nowadays, everything is digital, and there is a lot of e-commerce. So, not all e-commerce websites have efficient target URLs or web pages, which means that they can be attacked. They are open to attack right? So this is the major issue that I found in the phishing solutions. And because of that, there's a bias that fishing solutions can only detect. The attacks are based on popular fishing trends on these 30 intelligence platforms, like via social or phish tanks or open phishing. So so they are very much biased to the phishing trend and any new doc. Maybe there will be like, for example, there is alibaba.com is like very popular right? But there is like, for example, target.com is not very popular, but you can do e-commerce there, right? You can shop there was might be 2 or 3 fishing Urls associated with it. But not all.*". This introduces a bias into the efficacy of automated tools, thereby exposing users of less popular sites to heightened risks of phishing attacks.

> **Insight**: Rather than exclusively relying on automated anti-phishing solutions for anti-phishing defense, an increased emphasis needs to be placed on the employees' phishing education and training.

Phishers often impersonate reputable organizations, colleagues, or authoritative figures, making it more challenging for end-users to identify phishing attacks. In spear phishing, phishers incorporate content relevant to the users' interests, projects, or job roles to pique their curiosity. The difficulties end-users face while detecting spear phishing emails with personally relevant and familiarity are discussed in the literature, for example: "*because these messages mimic standard business processes, it can be hard for employees to tell the difference between malicious messages and safe ones apart*" [P58]. This data is also found in our interview study. During real phishing attacks in organizations, when attackers hack a legitimate email and continue sending emails to other employees by pretending to be an authorized person, it poses challenges to the employees to discern if the emails coming from this account should be trusted or not [U7].

Insight: As a precautionary measure, employees may treat emails as informational rather than instructional, given that - (i) a majority of phishing emails incorporate deceptive instructions, such as prompting users to click on links (ii) an attacker can send emails by pretending to be a legitimate contact.

The literature has discussed the challenges associated with ensuring the safety of employees operating within distributed organizational settings [P6, P54, P57, P62, P65]. Our investigation in this interview study revealed that, in numerous organizations, employees frequently engage in remote work or utilize personal devices where security measures are not deployed. Often, employees access unsafe websites using their organizational email credentials, thereby exposing themselves to the risk of phishing. A notable complication arises from this distributed work arrangement when this unsafe security behavior by employees leads to phishing incidents wherein attackers compromise email accounts. This is mainly because the timely dissemination of information to other employees about such phishing incidents becomes challenging, as not all employees routinely check their emails outside of regular business hours: "*For example the person's email, once it got hacked... happen on Friday. We got to know this on Monday*" [U6]. Typically, when employees utilize the organizational network, security teams undertake measures to isolate the employee's device from the network to forestall the spread of a potential phishing threat. Nevertheless, executing such actions encounters challenges in instances where employees operate on a different network. Thus, in such cases, there arises a necessity for an automated mechanism to promptly inform the IT security team of any suspicious activities associated with organizational email accounts.

Insight: • Develop a communication plan that outlines how to inform employees about the phishing incident • Use multiple communication channels to notify employees about the phishing incident • Provide clear guidance on what steps employees should take if they suspect any phishing attack.

Phishing attackers often leverage psychological tactics to manipulate end-users and distract them, increasing the likelihood of falling victim to phishing attacks. Distracted users often fail to critically evaluate the legitimacy of the content. Phishers take advantage of the fact that individuals may be fatigued or multitasking, causing them to be less vigilant. In such scenarios, users are more likely to overlook subtle indicators of a phishing attempt. Unfortunately, studies have shown that current intervention designs have not well-considered user distraction which affects the performance of the intervention outcome [P2, P7, P8, P13, P14, P24, P47].

This phenomenon is evident in our interview study, wherein we observed that certain organizations regularly send simulated phishing emails to their employees during active working hours to test their employees' susceptibility to phishing. When employees click on such phishing emails, they are redirected to a training page featuring guidelines on the appropriate and inappropriate actions in response to phishing attacks. Notably, Participant U3 explained that employees often ignore this training page when occupied with other work responsibilities.

Therefore, user distraction, workload, and working hours are significant elements current interventions should consider in the design and implementation. Despite receiving education about phishing, even well-informed employees occasionally succumb to the allure of clicking on phishing links when confronted with a substantial workload. The significance of distraction and working hours as a predominant factor in influencing the efficacy of phishing intervention outcomes has been discussed in Chapter 3.

> **Insight**: Organizations may schedule phishing training during times when employees are less likely to be heavily occupied with critical tasks.

The preparation of IT systems for phishing training presents a challenge, as documented in both the existing literature [P28] and our interview study. As previously mentioned, IT security teams send simulated phishing emails to assess employees' susceptibility to phishing attacks. These simulated emails incorporate contents that can trigger detection by automated anti-phishing solutions integrated into email clients. Consequently, the simulated phishing emails are often directed to the junk folder or categorized as spam, potentially missed by the employees. Addressing this issue necessitates security teams to manually configure rules facilitating the delivery of such emails to users' inboxes. Nevertheless, undertaking this task poses challenges for security teams, as underscored by the observations made by participant U2: "*So, if you use a very strict rule to remove those risks, it might lead to the fact that they [employees] missed some email or missed some information, so they might complain about that. However, if you relax a restriction, some security risks might get into our IT systems. So, that means our information security might be damaging. So that's why I said it is a challenge*" [U2]. In summary, security teams encounter challenges whereby the rules they generate may impede the delivery of routine emails to employees' inboxes.

> **Insight**: • Analyzing security event data can provide insights into emerging patterns and help refine security rules accordingly • Conduct periodic assessments of security measures and adopt adaptive adjustments to set rules to evaluate their effectiveness.

Literature has discussed the phenomenon wherein prolonged exposure to phishing warnings and training induces security fatigue among end-users [P4, P13, P14, P17, P18, P26]. This phenomenon is also supported by findings from our interview study [U4, U10]. Participants in our study mentioned instances wherein certain employees persistently clicked on phishing links. Although it is important to train these employees as often as they click on the phishing links, our participants emphasized that the primary objective of the security team is to avoid offending employees: "*...which is a difficult thing for security people, is that you cannot interrupt the users, you know. You'll have to work in the background... if you make the users unhappy, you know, they get annoyed*" [U4]. Nonetheless, it is necessary to train recurrent offenders, indicating a delicate balance that proves challenging to strike.

> **Insight**: • Organizations may consider shorter, more frequent sessions rather than long, infrequent ones to reduce employees' security fatigue. • Constructive feedback can be provided to repeat offenders to reinforce positive behaviors and correct their misconceptions. • Interactive workshops or group discussions can be hosted to share regular updates on emerging threats and best practices. • Rewards and recognition can be incorporated to motivate employees.

Several investigations in literature have addressed the phenomenon of end-users demonstrating a lack of awareness and attentiveness towards phishing warnings and training interventions [P1, P2, P3, P4, P8, P11, P14, P24, P28, P31, P36, P39, P44, P49]. The literature extensively outlines primary factors contributing to user disregard for phishing warnings, including the allure of visually appealing web content and the reputation of websites. Findings from our interviews corroborate the observation of employees exhibiting insufficient attention [U4, U6, U7, U12, U17, U15]. Our interview data reveal that employees engage in unsafe security practices, for example, they use weak passwords and visit potentially hazardous websites. Ignorance and inattention are pivotal elements that require consideration in the ongoing design of interventions, as also discussed in Chapter 3.

> **Insight**: The consequences of real-world phishing attacks can be well-communicated with the employees as a part of explaining to them the significance of adhering to robust security practices.

From the existing literature, it is evident that a considerable number of organizations lack compulsory phishing training programs for their employees.

Moreover, among those organizations that do implement such training initiatives, the execution is poorly organized. A significant proportion of these training sessions are conducted in an ad-hoc manner, lacking systematic testing or identification of specific groups susceptible to phishing attacks. These training sessions are typically led by an instructor who covers fundamental topics such as phishing, spear-phishing, and whaling. The attendance of employees in these training sessions is notably low. Consequently, a substantial portion of the workforce remains uninformed about phishing, leading to unchanged security behaviors: "*...About 30% of organizations favor the break room approach. They gather as many employees as they can in the break room, provide lunch, and have someone from IT or a security expert lecture on topics such as phishing, spear-phishing, and whaling. This is certainly better than nothing, but often attendance is low...And the results speak for themselves. Measures of the effectiveness of phishing show little change after such briefings*" [P59]. Literature also discussed users' unwillingness to install anti-phishing tools and software due to their complex installation procedures, which are sometimes not easily understood by non-expert users. These findings are similar to the findings of our interview study. Participants discussed the issues of lack of willingness among employees in their organizations to attend the training sessions: "*I would say the big challenge is not lack of interest, but lack of putting importance to this. So people attend this session just because they are told to, not because they're genuinely interested...They don't understand the consequences. That's why they don't care*" [U17].

Insight: • Phishing training methods need to be engaging and interactive to increase user participation. • Detailed instructional manuals need to be developed to explain the installation and operational procedures of anti-phishing tools and interventions to non-expert users.

Both the literature review and our interview study reveal the challenges faced by practitioners when engaging in the manual processes essential for planning, designing, deploying, and executing phishing training programs in organizations [P45, U5, U10, U15], as well as managing a substantial volume of phishing reports [P50, U1, U2]. Employees typically submit reports for suspected phishing emails, both during routine email interactions and simulated phishing training campaigns, resulting in a voluminous stream of reports that proves cumbersome to manage manually: "*Sudden large campaigns were found to overwhelm the help desk with reports, greatly impacting staff's workflow and hindering the effective application of mitigation and the potential for reflection*" [P50]. Our interview findings indicate that numerous organizations refrain from conducting phishing simulations due to resource constraints. Crafting phishing simulation emails necessitates substantial manual effort and time investment to ensure current and realistic content [P45]. Literature also discussed that phishing reports require manual human validation [P50]. Timely response to phishing reports is critical given the potential harm posed by phishing attacks within minutes [P50]. Additionally, security teams leverage these reports to enhance the defense mechanisms of their respective organizations [P50]. Proficiently handling

phishing reports demands expertise and meticulous attention. In addition to this, another challenge is the low rate of phishing reports as a lot of employees are reluctant to send the phishing reports: "*The first thing is nobody reported the phishing attack, the risks to us. So that means that we can't get timely and accurate risk information from our employees or our staff. Another ... a lot of people report a lot of emails or things to us*" [U2].

Insight: • Organizations may establish a ticketing system or centralized inbox to collect and organize phishing reports from employees and utilize email parsing tools to automatically extract relevant information (e.g., sender details, timestamps, and email content) from phishing reports. Then machine learning-based predefined rules can be set up to classify phishing reports based on severity or other criteria to prioritize responses. • An automated system can be set up to send acknowledgment emails to individuals who submit phishing reports to raise their motivation to report. • Where applicable reported incidents can be integrated into the training feedback loop to iteratively refine the system based on feedback.

Challenges to retaining user knowledge after phishing education and training is a concern discussed in the literature and also found in our interview observation [P13, P21, P40, P45, U4, U15]. Employees tend to forget the lessons learned on phishing: "*this is a tricky one, because there are tools in place, but then... they won't retain this information. The problem is, that the human mind cannot retain*" [U15]. Our investigation reveals that numerous organizations conduct induction-level training upon the onboarding of new employees or provide instructor-led optional training sessions, allowing employees the choice to participate (Section 4.3.1). In either case, organizations refrain from assessing users' comprehension of the information acquired during training. The data underscores the significance of probing users' enduring knowledge retention, a practice which is not commonly performed by the organizations studied in our sample.

Insight: • Organizations may send brief and recurrent reminders to their employees regarding the key principles of phishing defense.

The existing body of literature has identified a dearth of investigations into users' sustained behavior change [P7, P31, P34, P54]. Our conducted interviews provide substantiating evidence for these observations, revealing both confirmatory instances and challenges. Notably, as discussed in Section 4.3.1, certain organizations send their employees video-based training materials to educate them on phishing. Issues arise as employees frequently opt to bypass these instructional videos, leading to a lack of awareness and knowledge about phishing threats. The security team does not follow any formal approach to identifying whether an employee has viewed the video. Instead, they make assumptions grounded in specific informal criteria to understand the employees' engagement with the content: "*For example, if the person is completing it*

within 5 minutes, we know for sure that he just kept everything, if they implemented for 15min, we know like they had read something. If it's going for like 20 minutes they have not completed it. They have a threshold of timing based on that they just analyze this person like how they have read..." [U12]. A parallel challenge was articulated by our interviewee U6. Organizations that implement induction-based training programs often neglect to solicit feedback from employees regarding the effectiveness and comprehensibility of the training materials. Consequently, these unaddressed gaps in training often render employees susceptible to phishing attacks. Participant U16 expounded on the challenges associated with assessing users' enduring behavior change. Drawing from the experience in designing anti-phishing interventions, U16 highlighted the lack of a well-established framework specifying the set of questions that should be asked to users to test their knowledge of phishing.

Insight: • Organizations need to develop a formal procedure to track and monitor users' engagement with the training materials (e.g., whether the training videos are opened, clicked, or interacted with). • Security teams may collect user responses and feedback on the education and training materials to update these materials based on users' requirements.

As discussed in Section 4.3.1, numerous organizations conduct phishing simulation campaigns to assess the susceptibility of their employee to phishing attacks. In phishing simulation, a common approach for gauging employees' susceptibility involves quantifying the occurrences of employee clicks or the opening of malicious attachments. This challenge discussed in the literature, which supports our interview study findings, is the frequent miscounting of employee clicks attributed to the presence of bot clickers [P55, U18]. The technologies implemented within organizations often involve scanning incoming emails for potentially suspicious links or attachments. The presence of bot clicks in a phishing simulation campaign can complicate the evaluation of employees' susceptibility to phishing attacks. If the security system or simulation platform counts bot clicks along with legitimate user interactions, it may lead to inaccurate metrics and misinterpretation of the effectiveness of security measures. This miscounting can result in security teams erroneously believing that employees are interacting with phishing elements when, in reality, the interactions are automated: "*The other issue from a technological standpoint is what's called bot clickers. So if you have technology in your stack that tests emails coming in, some of them can explode the URL and test URL clicked. Those clicks also get counted as like people. The server response team is aware, and they follow the link, and then they explode it. So the person gets allocated training... They're like [employees], I didn't click on it*" [U18].

Insight: • Introducing human verification tests within the simulation (e.g., CAPTCHAs) can assist in identifying and filtering out automated interactions. • Leveraging bot detection mechanisms into a phishing simulation platform that analyzes user behavior, patterns, and characteristics can help to distinguish bots from human users.

4.4.1 Practitioners' perspectives on the guidelines

This section presents the findings of RQ5.3. Figure 4.5 and Figure 4.6 demonstrate the perceived strengths and weaknesses of the guidelines by the practitioners respectively.

TABLE 4.5. Practitioners' perspectives on the guidelines

Perspectives	Key points (included practitioners' ID)	#
① Guidelines convey meaningful information	• Guidelines are actionable and adaptable [U16] • Guidelines save time to look for resources [U6, U8] • Guidelines can be useful repository for implementers [U1, U12, U17] • Guidelines are reliable as they are collected from trustworthy sources [U11, U12, U13, U16] • Guidelines contain a good variety of instructions [U3] • Practical implementation would not be difficult [U12, U14] • Role based categorization is useful [U3, U14] • Rationale with the guidelines are helpful [U3, U14] • Guidelines are easily understandable [U14]	10
② Some factors necessitate consideration before adhering to the guidelines	• Prior investigation on the anti-phishing tool are conducted based on the budget [U10] (G14) • Providing system reliability information on the anti-phishing tools depends on the method used to design the tool [U14] (G14) • Information overload is a challenge while involving employees for feedback [U4, U15] (G28) • Organization conduct as many training cycles as possible as it never enough [U15] (G40) • Implementing progressive training is challenging as employees sometimes use mobile devices to read simulation emails [U4] (G35) • Implementation of the guidelines depends on the organization structure [U12] • Integrating guidelines to the existing security measures can be difficult due to policies and procedures [U10, U16] • Implementing Guidelines G40 and G41 would be challenging due to the intricate nature of achieving a balance between these guidelines [U10, U15] • Prioritising phishing warnings in the desktop and mobile devices depends on the application [U12] (G10)	6

Continued on next page

Practitioners' perspectives on the guidelines – continued from previous page

Challenges	Key points (included practitioners' ID)	#
③ Incorporate actionable instructions and required tools to implement the guidelines	• Include step-by-step instructions on the guidelines and provide real-world examples [U1, U4, U6, U7, U8, U9, U10, U13, U17] • Required tools to integrate the guidelines with the existing security measures need to be incorporated [U11, U15] • Need examples and info-graphics to improve understandability [U1, U5, U8, U9, U12]	13
④ Include statistical information on the guidelines to improve reliability	• More statistical information on the guidelines is required [U1, U4, U5] • Academic guidelines are sometimes not trustworthy as academics often do not make their data public [U4]	3
⑤ Adherence to specific guidelines necessitates more caution	• Employees sometimes ignore the embedded educational materials sent to them after phishing simulation [U4] (G12) • It is unusual for employees to skip the embedded educational material [U18] (G12) • If employees are pre-notified about phishing simulation their actual phishing susceptibility would not be revealed [U4] (G24) • Following a fixed email template may put organizations into more risks if the actual attack occurs [U7] (G26)	3
⑥ Guideline effectiveness depends on evolving phishing threats and organizations security culture and infrastructure	• Challenging to make employees adhere to the organizations' policy [U15] (G14) • Change the word victim mentioned in guideline G29 as it looks offensive [U4] (G29) • Guidelines effectiveness to resolve human-centric issues depends on the organization and its available tools and infrastructure [U4, U17] • Having an updated guideline is more important than a systematic guideline [U17] • To be effective the guidelines need to consider the constantly evolving phishing threats [U11] • Keep the guidelines updated and show comparison with other resources to make them more useful [U17] • Resolving the human-centric issues with the help of the guidelines depends on enforcing positive behavior to diverse range of users [U11] • Train data on guidelines on top of GPT models [U1]	5

Guidelines convey meaningful information

As demonstrated in Figure 4.5, most of our participants considered our guidelines reliable as our guidelines were devised from a systematic multi-vocal literature review: "*...I'm gonna say, absolutely I strongly agree that guidelines absolutely reliable...*" [U11]. "*I understand, you did a literature review. So I strongly agree that it came from this source, and I also think it is valid too, as a designer, because, we do a lot of user research and stuff. So I'm able to relate these kinds of things will improve that*" [U12]. As depicted in Figure 4.5, a majority (approximately 88% in Figure 4.5) of the participants think that

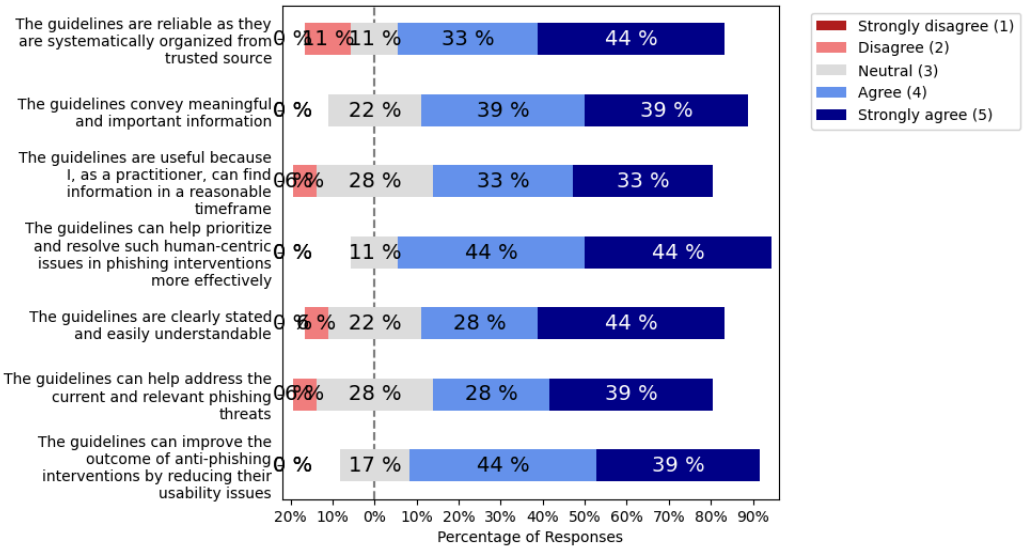

FIGURE 4.5. Perceived strengths of the guidelines

following our guidelines can help organizations prioritize and resolve human-centric issues in phishing interventions: "*...as far as I can see, the guidelines are pretty good like, because, you know, like most of the data breaches are the because of us humans because humans are the biggest link in cybersecurity...So this guideline is targeting*" [U16]. Some participants mentioned that the rationale provided with each guideline is helpful to make the guidelines more understandable: "*...I think it's really, you know, you provide more details [rationale behind the guidelines]. I think, in the table that you provided, there are more details into what that concretely looks like. So yeah, I think it's relatively easy to understand*" [U14]. "*The rationale, as I already mentioned. That's another strength*" [U3]. Participant U1 mentioned that our guidelines could assist the anti-phishing intervention designers in resolving the problem in the designs: "*It provides insights into the problems faced by designers and developers of user intervention phishing tools*" [U1]. Participant U1 also mentioned that our guidelines could serve as an initiative to develop a more useful repository on top of these guidelines: "*I think the strength of the guidelines is, it is a knowledgeable repository for finding more specific information*" [U1]

Insight: • The most appreciated attributes of the guidelines include role-based categorization, the incorporation of rationales underpinning the guidelines, and the compilation of guidelines derived from high-quality literature sources.

Some factors necessitate consideration before adhering to the guidelines

The integration and implementation of guidelines, alongside existing security measures, depend on multiple interrelated factors. The development of anti-phishing tools that offer system reliability information to the users, as recommended by our guideline G14, relies heavily on the methods and techniques

employed in their design. For instance, incorporating explainability into anti-phishing tools designed with deep learning models can be tricky due to their black-box nature. However, with the advancement of explainable machine learning domain, it would not be very difficult to add an extra block to these systems to generate information about the decision-making of the model: "*So I think it's [guideline G14] very specific to the machine learning or deep learning method that is being developed. But I think I'm aware of certain deep learning models that probably would have some difficulties, you know, to provide these kinds of explanations. Studies in the past several years have tried to make it more explainable. And we can actually add, you know, one additional block into the system that would provide more information about the decision-making*" [U14].

Collecting feedback from employees to update or modify organizational policies, as outlined in guideline G28, may result in information overload for users, leading to a tendency for users to disregard or inadequately engage in the feedback process: "*So again, it's one of those information overload like you have to go through all these things which we most of us don't care*" [U15].

Again, the assessment of anti-phishing tools before their adoption (guideline G25) depends on the organization's available budget and other prevailing priorities: "*well, the first major thing that comes into it is budget. So, like, if we have the budget for it. Which kind of sucks, but then we do conduct options paper. So then we'll nail it down to well, let's just say 4 options. And then after that, we'll do an extensive analysis on those*" [U10].

Implementing progressive training (guideline G35) presents challenges, given that employees often use mobile devices to access simulation emails while traveling. Our guideline G10 suggest designing a uniform phishing indicator across desktop and mobile device to reduce the risk of mobile users. Participant U12 suggests that designers and developers prioritize certain things based on the application they are developing, also the target audience of that application. Developers do not usually prioritize security over other requirements. Participant U12 underscores the challenge in uniformly prioritizing all requirements, attributing this difficulty to variations in screen dimensions: "*So if I'm developing a product if it's a bank Financial product the main important aspect is the user needs to trust the product. So if that's the case, we will put security as the top priority and the next one as the second priority. If it's just a kid, ... kids need not know...if the app is secure or not, because they don't know anything about that. They just need to play the games, they just need to color the book on the app which is in the application. So they don't care about the security aspects*" [U12].

Certain participants provided a neutral response regarding the feasibility of incorporating our guidelines into their organizational security framework as shown in Figure 4.6. We observed that the incorporation of our guidelines depends on factors such as organizational size, the presence of available tools, the time associated with establishing new rules and protocols, and the availability of employees capable of effecting the recommended changes outlined in the guidelines [U10, U12, U16].

Some participants [U10, U12, U15] in our study highlighted that determining the optimal number of training cycles to effectively educate and train employees

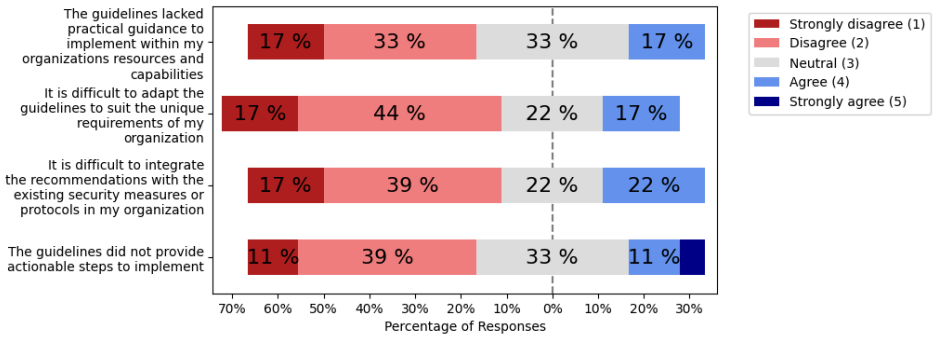

FIGURE 4.6. Perceived weaknesses of the guidelines

remains an ongoing challenge that warrants further investigation. The majority of respondents expressed concern about avoiding employee fatigue due to excessive reminders, emphasizing the importance of training them according to their actual needs.

Insight: • Our observation indicates an area that requires further investigation is how organizations can ascertain the optimal number of cycles of training for recurrent offenders without inducing security fatigue. • To integrate any guidelines in the organization, it is important to know the deployed security measures, policies and procedures, and budget constraints of the organization.

Incorporate actionable instructions and required tools with examples and illustrations

To understand the weaknesses of the guidelines, we asked the participants about the practicability and adaptability of the guidelines. As shown in Figure 4.6, most of our participants think that our guidelines can be adaptable to suit the requirements of their organizations and our guidelines can be integrated into the existing security measures and protocols of their organizations. However, some participants shared insights to improve the practicability of the guidelines. According to some participants, the guidelines would be more useful if step-by-step guidance for each guideline could be incorporated. It is suggested to add implementation details on the guidelines, for example, possible tools and infrastructures required to implement the guidelines in the organizations. This is mainly because implementing the guidelines necessitates formal policy-making and tools being made available in the organization. Again, a common suggestion to improve the guidelines' practicability was to include examples with the guidelines: "*So my point is making it specific with examples*" [U6]. "*...give an example so that people can understand better...*" [U1]. Our participant U6 agreed that some guidelines are actionable and mentioned some guidelines are very broad which requires practical examples: "*if it's very specific, then you don't need to give an example... So if it's broad and generic, then an example would help*" [U6]. Our participant U9 is a CEO of a Small and Medium Enterprise (SME) who had difficulty due to a lack of security and technical

knowledge: "*if you want my sort of opinion on it, like as a C suite employee, and also as somebody who you know, is not a technical person. I find it difficult to understand*" [U9]. Therefore, to improve the guidelines' understandability to the C-suite employees who are not security experts but are involved in decision-making, the use of infographics and explanatory notes with the guidelines can improve their understandability. To demonstrate the usefulness of the guidelines, practitioner P1 suggested highlighting the difference between guidelines and the answers generated by large language models such as ChatGPT. How much the guidelines are capable of catering to the needs of practitioners compared to other available resources such as ChatGPT was also suggested to include with the guidelines. To help organizations integrate the guidelines with their existing security measure, participants U11 and U15 suggested adding the tools required to implement the guidelines within the organizations: "*from a customer perspective, I think what you have, they will be good. The only thing is, normally, if we look at that from a white perspective, which tools do we have to do that?*" [U11].

Insight: • The implementation of the guidelines within organizational contexts necessitates the provision of actionable details, substantiated by real-world examples that incorporate relevant tools and technologies.

Include statistical information on the guidelines

From Figure 4.6, it is noticeable that some participants provided neutral or negative responses as they wanted to have more information on the guidelines. Industry practitioners often do not follow or trust the guidelines or suggestions provided in academic literature. This is mainly because sometimes, in academic studies, authors do not disclose detailed information about their data and methods. Therefore to increase the credibility and reliability of the guidelines our participants U1, U4, and U5 suggested adding detailed descriptions of when the data was collected (recent or outdated), which industries were considered in the primary studies, what statistical and demographic information was mentioned in the studies that were used to devise the guidelines etc. "*when was this data collected? Is it recent, or old? Or what sort of industry that is talking about? what is the population?*" [U4]. When we devised the guidelines systematically from the recommendations provided in the academic and grey literature, we did not summarize the demographic data considered in these studies, such as the number of participants considered in these studies, any background details explanation of how those studies have come up with these recommendations, etc. Participant U5 suggested summarizing these demographic details along with each guideline to make the guidelines more reliable: "*when we were talking about narrative-based training, you were telling me, like research says that this kind of training is more suitable to this demographic of people... between these ages... So that kind of information is more useful for me*" [U5].

Insight: • Incorporating contextual elements and an evidentiary foundation (e.g., dataset utilized, demographic characteristics of the studied population, contextual information on the study settings, any inherent assumptions made during the investigation) enhances the credibility of the guidelines for practitioners.

Adhering to specific guidelines necessitates more caution

In this section, we examine participants' concerns and insights on our guidelines G12, G24, and G26. According to guideline G12, in phishing simulation and embedded training, interventions (specifically landing pages) should be promptly presented to users immediately after they click on any simulated phishing links, aiming to provide instantaneous education to employees. Participant U4 contributed to the discussion, expressing that landing pages could be seamlessly integrated into the training. However, the suggestion of dispatching video-based training materials after employees have clicked on simulated phishing links was deemed impractical. This stems from the concern that such materials might be ignored by the employees, particularly if sent during business hours: "*But the problem is like, if someone clicks and they're in the middle of the work, they don't have time to watch a video or spend a couple of minutes, you know. So what do we do? We have a one pager [landing page] that pops up there, and then, you know, it shows that why they failed it. But we cannot do much more than that, because you know, users are busy with work, and we cannot just say we have to do the training right now. Okay, that's the challenge. And also like, if you don't have statistics about what sort of things users are missing, you cannot customize your program*" [U4]. It is noteworthy that participant U4 may have found our guidelines less clear, as guideline G12 primarily refers to landing pages rather than video-based training materials. Participant U18 offered a contrasting perspective, asserting that the likelihood of employees neglecting the landing page is minimal, as it is presented within the browser on the same window: "*But the landing page generally if you click the link... it'll open up in the browser right in front of you. So, except for the people who close it before it's open properly... Most people will see it. So I don't really know about people skipping it. I don't really understand how ...*" [U18].

Guideline G24 suggested sending pre-notification to the employees before conducting the phishing simulation in the organization. According to practitioner P4, this action may make some employees alert in advance which can give a false impression of employees' security behavior and actual phishing susceptibility: "*But if you tell them we're doing a phishing campaign now, game over. But then some companies are doing like... tell them a month ago. So is okay*" [U4]. In our guideline G24, emphasis was placed on the implementation of randomized phishing training, wherein employees are made aware of the occurrence of a simulation but remain uninformed about its specific timing. It has come to our attention that Participant U4 found this aspect of G24 less lucid, as reflected in the above statement, which also underscores the importance of random simulations. A more precise articulation of the language may enhance the clarity of G24.

Our guideline G26 suggests organizations have a fixed email template to help employees identify the irregularities in the phishing email. However, participant U7 shared an experience of dealing with a real phishing attack in his organization where attackers, after compromising the email, remained silent and observed other employees' behavior and email communications for a few days to receive trustworthiness from those employees. Therefore, according to participant U7, following a fixed email template can be counterintuitive: "*So what happened that when ...they [attackers] infiltrate one of their [Organizations'] emails they do not instantly change the password and take control of that email address instantly rather than they actually keep the presence inside the email and check the transactions for months. And they understand. And they try to get to know about the structure of the email. And when they learn enough about this structure, they start sending those [phishing] emails*" [U7].

Insight: • Guideline G12 can rephrased as: "*During a phishing simulation campaign, display the landing page promptly to users after they click on any simulated phishing links*" • Guideline G24 can be re-phrased as: "*Send a pre-notification to users well in advance of conducting phishing simulation training, withholding the precise date of the simulation*" • Guideline G26 can be re-phrased as: "*Adhere to a standardized template for organizational email communications, while explicitly advice employees not to unquestioningly trust any incoming email conforming to the prescribed template. Create a standard template for anti-phishing webpages*"

Guideline effectiveness depends on evolving phishing threats and organizations' security culture and infrastructure

Approximately 28% of study participants provided neutral responses to the statement concerning the efficacy of the guidelines in addressing current phishing threats, as illustrated in Figure 4.6. Furthermore, Figure 4.5 reveals that approximately 11% of participants expressed neutrality regarding the guidelines' effectiveness in prioritizing and resolving human-centric issues in phishing interventions. This response is attributed to the dynamic nature of phishing threats, necessitating regular updates on the guidelines [U4, U11, U17]. Additionally, the effectiveness of the guidelines in addressing human-centric issues depends on the organization's security culture and enforced policy [U15]. Participant U4 recommended revising the term "victim" employed in guidelines G29. Participant U1 highlighted that our static guidelines are inadequate for fulfilling organizations' unique requirements. A suggestion was made to augment these guidelines with training data integrated into GPT models to accommodate more distinctive organizational needs.

Insight: • For optimal efficacy, guidelines necessitate periodic updates in response to the dynamic evolution of phishing threats, attack methodologies, and defensive measures • Guideline G29 can be rephrased as: "*Implement help-desk assistance services for users requiring support following a phishing attack*"

4.4.2 Desired features of an envisioned tool to access guidelines

This section presents the findings of RQ5.4. Figure 4.7 and Figure 4.8 display the perceived best aspects and limitations of the tool *PhishGuide* by the participants respectively. Both Figure 4.7 and Figure 4.8 demonstrate that the majority of our participants provide positive feedback on the tool prototype *PhishGuide*. Figure 4.7 highlights that most of the participants appreciated the organization of the options in the tool, the categorizations provided, the number of steps required to get a guideline, the personalization provided by the tool, and our effort to organize the guidelines for the practitioners to save their time. Again, Figure 4.8 depicts that according to most participants, the tool is easy to use and easy to learn and the amount of information provided in the tool is reasonable and not too much to read. A small number of participants who provided neutral responses and negative responses suggested some improvements to the tool. We will discuss these suggested improvements in the next sections. Anti-phishing tool developers and researchers can take inspiration from these suggestions to design tools for security practitioners in organizations.

Incorporate customizable tool options

While practitioners appreciated our prototype tool *PhishGuide*'s guideline generation based on intervention stages and practitioner groups, it doesn't suit those who serve multiple roles. For instance, practitioner P6 is both a developer and involved in implementation. Similarly, practitioner P5 handles implementation/evaluation but also designs training content. The tool should have design options for security team members and needs implementation/evaluation options for anti-phishing intervention developers. Again, the tool should offer visible options alongside personalized ones to prevent practitioners from being uncertain about what they might have missed.

***Key points*:** • Personalize tool options based on practitioners knowledge [U9, U17] • Role based categorization is interesting [U14] but can be misleading [U6] • Add design option for security team members [U5] • Along with personalization, display other options [U17] • Tailored options save time [U12] • Minimize the options required to get a guideline [U4].

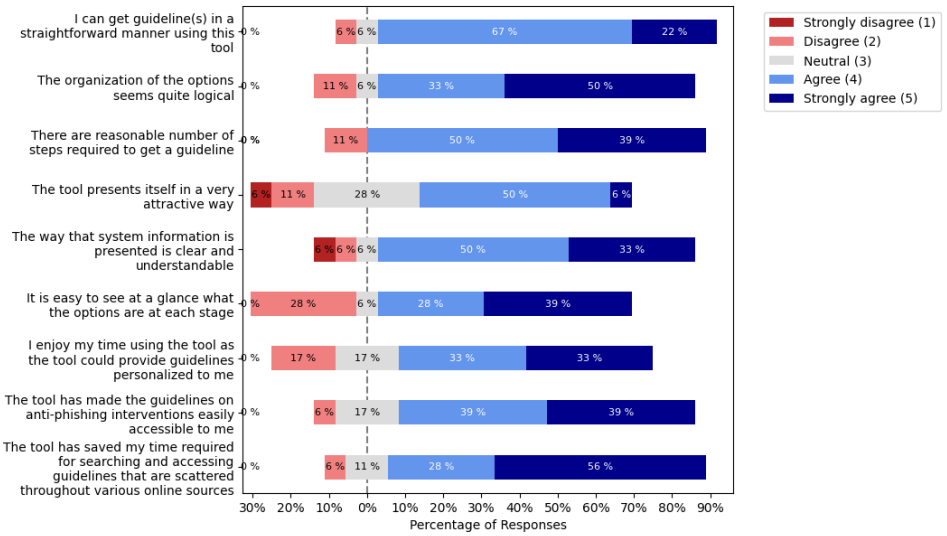

FIGURE 4.7. Perceived best aspects of the tool *PhishGuide*

Integrate responsive features and real-time feedback mechanism

Practitioners expressed a preference for the integration of a search bar within the tool, enabling them to input their specific requirements for accessing guidelines and searching for definitions of various terms. The implementation of clickable links and a drop-down menu was recommended to maintain a user-friendly interface, consolidating all features within a single window to minimize time and effort. Additionally, it was emphasized that the tool should allow practitioners to obtain guidelines without necessitating a return to the initial starting point, ensuring continuous flexibility in navigating between options. Furthermore, the incorporation of a progression bar was proposed to indicate the users' progress toward obtaining a guideline.

> ***Key points***: • Add search bar to search for tool documents and definitions [U1, U3] • Add a progression bar to display the remaining steps [U2, U8, U17] • Allow users to navigate through different options without having to go back to the starting point [U13, U16] • Add Natural Language Processing (NLP) based search query to generate a guideline [U1, U12] • Add clickable links, drop-down menu and keep everything in one window [U4, U9, U11, U16, U17] • Introduce automatic update of new contents [U10, U15].

Provide demonstration on tool features

Practitioners express the need for comprehensive elucidation concerning the terminology employed within the tool, straightforward delineation of the tools' functionality or capabilities through visual aids such as videos or images, and clear distinctions among diverse categories. In essence, the tool should furnish adequate information for novice employees and those lacking technical expertise,

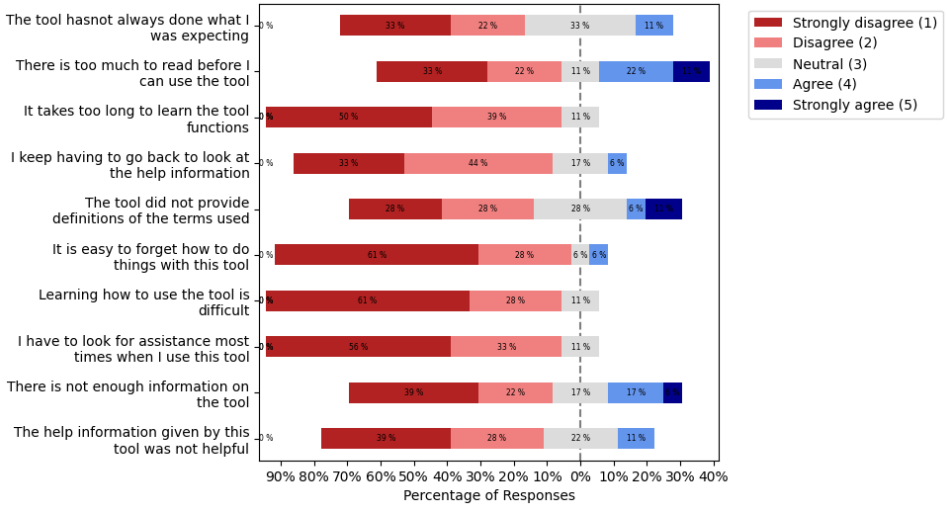

FIGURE 4.8. Perceived limitations of the tool *PhishGuide*

obviating the necessity for extensive background knowledge to effectively utilize the tool.

Key points: • Provide a brief summary of the tool functionalities and capabilities [U4, U5, U8, U17] • Use website or mobile app to commercialize the tool and add a table of contents [U4, U8, U9] • Provide definitions of the terms used in the tool [U5, U8] • Summarize the importance of the tool [U11] • Add a small clip for the new users to describe the tool functionalities [U6].

4.5 Recommendations for Researchers and Organizations

In this section, we discuss the implications produced for the researchers and organizations based on our findings.

4.5.1 Investigation on the challenges to integrating academic recommendations into organizational practices

Numerous academic studies have offered recommendations to enhance anti-phishing practices within organizations (e.g., [12], [162], [136]). However, the findings of the evaluation of our guidelines [139] discussed in this chapter in Section 4.4.1 reveal that real-world implementation of these guidelines requires consideration of several factors. Therefore, simply providing recommendations to organizations is insufficient in effectively countering phishing attacks. Addressing the actual challenges faced during the implementation of these guidelines in

real-world scenarios is crucial. This assertion is supported by a recent study that argued that industry practices do not align with cutting-edge human-centered security approaches [163]. For instance, our guideline G31 in discussed in Chapter 3 [139] proposed integrating embedded training with phishing simulation to educate employees within an organization. Contrarily, a recent study [22] involving employees from various organizations demonstrated that embedded training might heighten employees' susceptibility to phishing attacks. Similarly, the adoption of phishing simulations to assess employees' vulnerability to phishing attacks is also advocated in the literature. However, a recent case study [66] highlighted the concealed expenses associated with the selection, procurement, and implementation of phishing simulations, rendering it unfeasible for many organizations to adopt. Hence, it is imperative to conduct further research on the feasibility of academic findings concerning the design, implementation, and evaluation of anti-phishing interventions in practical organizational contexts. Merely proposing recommendations based on theoretical frameworks proves inadequate; understanding the practical challenges faced during their application is indispensable for developing effective anti-phishing strategies.

4.5.2 Tailored training methods for organizations

Our findings discussed in Section 4.3.1 show that our studied organizations use diverse types of education and training methods to educate and train their employees, for instance, video-based training, instructor-based training, email-instruction based education, phishing simulation, and embedded training. From the interviews, we observe that the organizations are adopting these education and training methods without knowing or properly investigating which method would be suitable for their employees: "*...have you got any evidence from research saying which [training method] has been more effective?*" [U5]. To date, different education and training approaches have been proposed, discussed, and evaluated in the literature. For example, quiz and online seminar-based training [164], persuasion principle-based training [165], game-based training [50], facts, advice and story-based training [39], embedded phishing training [19] etc. Unfortunately, to the best of our knowledge, no existing study has rigorously investigated the strengths and weaknesses of different education and training approaches. This investigation is necessary to provide organizations with guidance on selecting an appropriate method tailored to their unique organizational settings, and requirements. Some of the unique requirements can be workforce availability, financial constraints, organization size, the availability of tools and infrastructure, etc. Tailoring phishing training to the organization would be beneficial as suggested by a recent study [63].

4.5.3 Feedback from employees and follow-up on phishing education and training

The organizations in our study provide education and training to the employees either by sending employees video-based training materials or text-based instructions or providing instructions in a seminar conducted by experts. None

of the organizations have appropriate formal procedures in place to receive feedback from employees on the training materials and methods. It is unknown if the training methods were useful, enjoyable, and understandable to the employees. In this way, organizations may spend dollars on buying phishing simulation programs (phishing-as-a-service) [66] but these would not change employees' security behavior until employees' feedback is incorporated. A recent study also reported that organizations are very dependent on what phishing vendors offer and human constraints and employee requirements are not considered [163]. Based on the aforementioned discussion, we suggest the integration of employees' suggestions and requirements in the training materials and methods to make them engaging, interesting, and understandable.

4.5.4 Guidance to design training content with real-world examples

In literature, the efficacy of employing facts, advice, and narrative-driven phishing training methods has been substantiated as effective approaches for enhancing individuals' retention of anti-phishing instructions [39]. These techniques have also been acknowledged in our study: "*But when you actually come up with different practical scenarios that already had happened and then relate those content with those scenarios people listen, and people understand. And people remember those things*" [U7]. Unfortunately, our findings discussed in Section 4.3.1 and Section 4.3.1 reveal that practitioners often encounter challenges in accessing structured guidelines and a sufficient number of practical instances to integrate into their training materials. Although there are numerous phishing URL datasets and platforms (e.g., PhishTank [166]), we are not aware of any resources or platforms containing genuine phishing narratives, news accounts, and their consequential impacts on individuals, data, and information systems. Consequently, it is recommended that a comprehensive database of phishing incidents be established, providing online users with a valuable resource to learn about and comprehend the intricacies of phishing attacks.

4.6 Limitations of the study

Notwithstanding the difficulty in recruiting industry practitioners for this study due to their high workload and limited contactability by outsiders [145], we managed to interview 18 practitioners of diverse roles and experiences who fulfilled our inclusion criteria. Our final sample comprises 18 practitioners from 18 different organizations of various sizes and diverse domains from 6 countries. We acknowledge that a majority of the data is gathered from organizations situated within Australia, totaling 12 diverse organizations. This outcome may have been influenced by our social media connections and followers, as our recruitment strategy entailed posting advertisements on our social media accounts. Although we sent invitations to the practitioners across various countries, the decision to participate in our study ultimately rested with the practitioners themselves.

For the current settings, we have reached theoretical saturation as no new codes were produced after a certain number of practitioners. However, qualitative studies are inherently interpretive, and the findings are based on the studied context, thus challenging to generalize [167]. Therefore, future studies can further extend or evaluate our study findings by following a different recruitment strategy and study settings.

The desired tool features we extracted from the participants' discussion might have been impacted or influenced by the features of tool prototype *PhishGuide* we presented to the practitioners during the interview. To mitigate this threat, besides having close-ended questions on the tool, we asked the practitioners what features would they expect from such kind of tool aimed at generating guidelines.

To enhance internal validity, stringent inclusion criteria were employed, limiting the recruitment to practitioners engaged in the design, implementation, and evaluation of anti-phishing tools, interventions, and technologies. However, given the diverse nature of the guidelines, some of which were tailored for C-suite employees, participants not directly engaged in anti-phishing interventions were also recruited. This inclusion was warranted as these practitioners play pivotal roles in the organizational decision-making process.

To address the potential threat to construct validity, the first author developed the interview guide, which was subsequently reviewed by the remaining authors. Additionally, a pilot study was conducted by following the interview protocol to test the study settings.

We developed closed-ended questions comprising multiple statements in both the guidelines and the tool, aiming to gain insight into practitioners' challenges, requirements, and preferences. During the interview, we asked participants to explain their opinions on each statement while answering the questions. This approach facilitated the capture of nuanced details, thereby mitigating the risk of potential misinterpretations in the collected data.

None of the participants in the study were engaged in the development of phishing warnings, including email client phishing indicators, browser warnings, and browser anti-phishing toolbars. Consequently, in order to enhance the comprehensiveness and accuracy of the data collected, we encourage future investigations to involve surveys, interviews, and case studies with developers responsible for creating phishing warnings.

4.7 Chapter Summary

In this chapter (i) we investigated the existing anti-phishing mechanisms implemented within organizations, aimed at safeguarding their people, data, and infrastructure against phishing attacks (RQ5.1); (ii) we reported 8 challenges in practice in the design, implementation, and evaluation of anti-phishing interventions (RQ5.2); (iii) we evaluated the 41 personalized guidelines presented in Chapter 4 for the anti-phishing practitioners to support better design, implementation, and evaluation of 14 different interventions (RQ5.3); (iv) we collected desired features from the practitioners for an envisioned tool aiming to support developers and practitioners to access the guidelines (RQ5.4).

The findings we reported in this chapter have derived an evidence-based understanding of the challenges and requirements of anti-phishing practitioners involved in the design, implementation, and evaluation of anti-phishing interventions, which have several implications for researchers and cyber security practitioners. Our results on anti-phishing practices (RQ5.1) provide opportunities for cyber security experts and decision-makers to modify and update existing security policies and protocols of organizations to protect themselves against phishing. Also, these findings on anti-phishing practices can potentially guide cyber security experts to impose new rules if necessary or to take the right initiatives that organizations currently lack to protect them from phishing attacks. The challenges documented in our study (RQ5.2) encompass issues such as limitations in anti-phishing datasets and tools, constraints related to organizational resources, and challenges in designing training content. These challenges have created an opportunity for researchers to delve deeper into this area and devise improved phishing education, training, and awareness mechanisms to counter phishing attacks. Our findings on the evaluation of the guidelines (RQ5.3) in the literature aim to offer valuable insights into the formulation of guidelines to improve the design, implementation, and evaluation of phishing education, training, and awareness interventions, emphasizing their comprehensibility and usefulness in practice. Furthermore, we explore the obstacles encountered during the practical implementation of these guidelines and report recommendations by the practitioners aimed at enhancing their effectiveness. Our reported actionable tool features (RQ5.4) can guide future phishing researchers and anti-phishing tool developers in this domain to design usable tools to make academic research findings easily accessible to anti-phishing technology developers and security practitioners.

Chapter 5

Conclusion and Future Research Directions

5.1 Summary of the Contributions and Findings

This section summarizes the contributions and findings of this thesis. This thesis contributes to the existing literature in eight substantial ways, as detailed below.

5.1.1 Systematizing the socio-technical challenges reported in the literature in the design, implementation, and evaluation of phishing intervention

We present a pioneering systematization of knowledge on challenges in phishing interventions, encompassing 8 design challenges, 7 implementation challenges, and 5 evaluation challenges (Chapter 2). The identified 8 design challenges (Section 2.5.1) are UI design restrictions in the browser and email client, content restrictions for phishing education and training, design constraints for anti-phishing warning UI interfaces, problems with anti-phishing warning content, performance limitations of anti-phishing tools, lack of attention to phishing indicators, need to design specific training for spear phishing, disregard for users' mental limitations during design. The implementation challenges (Section 2.5.2) are anti-phishing technology deployment challenge, technology adoption and usage challenges, challenges due to complicated URL and domain name structures, obstacles to automate phishing incident response and anti-phishing training, exploitation of software vulnerabilities by attackers, unguarded email clients and websites and limitations of current anti-phishing planning, policies, and guidelines. The evaluation challenges (Section 2.5.3) are lack of industrial relevance in evaluation practices and settings, complications regarding data collection and replicating user experience, insufficient usability and effectiveness evaluation of phishing interventions, lack of sophisticated quantification of phishing training outcome and lack of post-training user knowledge retention practice.

Our analysis underscores numerous requirements that are currently absent or inadequately considered in the phases of design, implementation, and evaluation. Notably, we identify deficiencies such as the absence of active interruption

in phishing warnings and suboptimal warning placement, UI design inconsistency of phishing warnings for mobile and desktop browsers, absence of phishing indicators in the email client for forged emails, complex user interfaces of phishing education games, lengthy and wordy phishing training contents, lack of comprehension in the phishing warning contents, lack of consideration of human-centric limitations in the design of phishing warnings and so on. Additionally, we shed light on implementation challenges arising from browser platform dependencies and the complexities introduced by distributed work settings and extended infrastructure within organizations, complicated domain names/URLs, lack of use of SMTP extensions (such as Sender Policy Framework, Domain Key Identified Mail) for email client authentications, contradictory and abstract anti-phishing recommendations in the organization's website, etc. Our investigation uncovers usability issues inherent in phishing interventions, stemming from an oversight in evaluating the diverse demographic features of end-users during the early stages of prototype construction for these interventions. We also identify that organizations' phishing simulation and training outcomes are sometimes misinterpreted due to the presence of the bots from the security tools and third-party bots. In summary, our research insights provide practitioners with a comprehensive understanding of the constraints associated with anti-phishing interventions.

5.1.2 Systematizing the critical success factors in the design, implementation, and evaluation of phishing interventions

We undertake to methodically categorize the disparate recommendations to provide a comprehensive view of factors that contribute to the efficacy of phishing interventions. We synthesize 23 success factors concerning the design, implementation, and evaluation of phishing interventions (Chapter 2). We extracted recommendations on designing more interesting and engaging phishing education and training content design (adopt gamification, include video materials, include less text, add more graphics, etc.) to strengthen user motivation and increase content consumption, the creation of diversified, current, and captivating training content, the incorporation of dynamic and self-adaptive training methodologies, the development of unified phishing indicators applicable across diverse browser platforms and devices, designing explainable anti-phishing systems and technologies (e.g., creating user-friendly URL patterns, providing explainable reports for automated anti-phishing solutions) to help users gain a better understanding, tailoring the design of anti-phishing interventions (e.g., provide individualized training contents for the children), improve the UI designs of the phishing interventions (e.g., adding a support button in the email client phishing warning for non-expert users, employ varying text sizes and colors).

Our findings also highlight the importance of the participation of different key stakeholders to encourage the end users and employees of their organization, and the importance of strong authentication and encryption mechanisms for incoming emails handled by email clients. We also found from the literature

that it is necessary to arrange follow-up training to reinforce user knowledge of phishing, to conduct GDPR compliment and anonymous training in the organizations to protect user's privacy, and to provide special attention to critical demographic groups. Our results shed light on the importance of automating the implementation tasks (e.g., handling the phishing reports) to assist organizations' security teams and emphasize better planning and policy management on phishing training. In short, our investigation yields original insights designed to enhance the effectiveness of anti-phishing initiatives within complex real-world environments.

5.1.3 Investigation of the role of socio-technical factors in influencing the effectiveness or susceptibility to phishing

We have identified a novel set of 22 socio-technical factors, including 15 human factors, 4 technical factors, and 3 organizational factors essential for customizing the phishing interventions (Chapter 3). Our conducted study reported in Chapter 3 reveal that several demographic factors (e.g., age) and user's cognitive constraints (e.g., knowledge decay) are currently not considered or ill-considered in the design of phishing intervention (Section 3.3). For example, research has shown that children, teenagers, and old people have different requirements for phishing training due to their knowledge level or how they respond to phishing emails. Again, educational qualification impacts the outcome of phishing training in different ways, for college students phishing stories from a peer is an effective method of training whereas for university staff, an expert-based phishing training method is a more suitable approach. Again, an individual's personality also affects the outcome of the phishing training, for example, overconfident users over-trust their ability to detect phishing emails causing them to disregard the phishing emails. We also observe that organizational factors such as subjective norms and social influence play a vital role in end-users ability to detect phishing attacks. In summary, our investigation underscores the necessity of taking into account various end-user demographics, including age, educational qualification, knowledge level, and organizational position, in the tailoring of interventions.

5.1.4 Customized guidelines for four practitioner groups involved in the design, implementation, and evaluation of phishing interventions

We present 41 tailored guidelines on the design, implementation, and evaluation of 14 distinct categories of anti-phishing interventions across four professional cohorts: designers/developers, information security team members within organizations, cyber security experts, and executive-level personnel in organizations (Chapter 3). These guidelines are methodically formulated in response to our identification of 23 pivotal success factors, aimed at addressing 19 recognized challenges and integrating 22 socio-technical considerations. Our guidelines

provide numerous design recommendations to the developers and information security team members on the user interface design of the interventions, phishing education and training content design, improvement of the decision-making feature of anti-phishing technology, different ways to personalize the style and medium of the intervention and so on. We report the implementation guidelines that include the discussion on the configuration of the IT systems for phishing training, improving the log-in page of a website, suggestions on the adoption of anti-phishing tools for the organizations, use of templates for anti-phishing webpages, guidelines on organization's policy improvement, etc. Our guidelines include recommendations for improving the evaluation process of anti-phishing interventions such as choosing the evaluation metrics and baselines. The guidelines we have developed serve as a valuable reference for practitioners seeking to enhance their ability to combat socio-technical challenges inherent in the formulation, execution, and evaluation of anti-phishing interventions.

5.1.5 Systematic overview of anti-phishing measures implemented in organizational settings

We present a systematic overview of anti-phishing practices implemented within organizations (Chapter 4). This empirical analysis encompasses the methodologies and content associated with phishing training, the frequency of training sessions, the procedures for reminders and notifications, the assessment processes applied to evaluate employees' knowledge, and the manual and automated mechanisms employed for phishing detection within organizational contexts. Additionally, it encompasses the responsive measures taken in the event of an actual phishing attack.

We observe that, our studied organizations employ a range of training methods including video-based induction level training, phishing simulation, and embedded training, instructor-based training to educate and train their employees on phishing attacks. For phishing training content design, we notice that the organization's security officers do not have enough resources and support to include real-world phishing examples in the training content and to personalize the content. A majority of our studied organizations do not conduct any follow-up training whereas some of them run multiple cycles regularly. After conducting phishing training, most of the organizations do not asses the phishing knowledge of the employees to test their knowledge retention. In summary, our reported body of knowledge on anti-phishing practices in organizations can serve as a valuable resource for security experts and decision-makers, offering insights that can guide the implementation of required changes to fortify organizations against phishing attacks.

5.1.6 Identification of challenges in practice to safeguard organizations against phishing

We acquire an empirical understanding of the extant challenges inherent in the design, implementation, and evaluation processes and discern a total of nine challenges in practical application (Chapter 4). Subsequently, we analyze these

nine challenges, ascertained through empirical observations, with nineteen challenges documented in the literature (discussed Chapter 2). This comparative analysis seeks to elucidate both commonalities and novel insights. Our empirical examination illuminates tangible challenges existing in real-world scenarios, including impediments in the design of training content, constraints associated with anti-phishing datasets, limitations of training materials, and challenges related to instilling motivation among employees to promote secure behavioral practices. The exploration of these practical impediments facilitates a nuanced understanding of how our formulated guidelines have the potential to address these challenges effectively.

5.1.7 Understanding the barriers in implementing the devised guidelines in practice

We present findings derived from industry practitioners to assess the efficacy of our formulated guidelines and to comprehend the practical challenges inherent in their real-world implementation (Chapter 4). Our approach involves the systematic collection of multiple recommendations on our guidelines, and we subsequently elucidate the ramifications of adhering to these guidelines within diverse organizational contexts and settings. Our investigation with 17 practitioners from 18 organizations from 6 different countries discusses several perceived advantages and limitations of our guidelines. Participants of our study (reported in Chapter 4) find our wide variety of guidelines meaningful and understandable. They mentioned that our guidelines can be a useful repository for them as they are derived from trusted sources. We also discussed the recommendations provided by the participants to improve our guidelines such as inclusion of visual examples, real-world implementation details, and relevant statistical information (if any) for the guidelines. Our interview study reveals that some issues require in-depth investigation such as identifying the suitable training cycles for the organization and finding budget-friendly anti-phishing solutions for small and medium-scale organizations. Through our empirical analysis, we gain insights into the utility and applicability of the guidelines we have devised.

5.1.8 Compilation of features for a prospective tool facilitating practitioners' access to guidelines

We provide a consolidated overview of functional attributes and prerequisites essential for a prospective tool, which can enable practitioners to access our guidelines (Chapter 4). These delineated features of the tool serve as guidance for researchers and tool developers in formulating instruments that integrate functionalities, including but not limited to, natural language processing-based search queries for the generation of guidelines and the automatic updating of new guidelines. We also gathered desired functional features on the tool such as adding a search bar to allow for searching relevant documents and non-functional features such as adding a small video clip for the users who are new to the system.

5.2 Opportunities for Future Research

To the best of our knowledge, this thesis represents the first effort in the literature to enhance the design, implementation, and evaluation of phishing education, training, and awareness interventions by synthesizing challenges and critical success factors compiled from academic and grey literature, factors to tailor the interventions, personalized guidelines to support practitioners, and documenting the challenges and requirements of practitioners in this area. Although we have reported several recommendations in Chapter 2, 3 and 4 based on the three studies conducted, the knowledge can be further extended through replication studies, re-evaluation, considering large-scale settings and designing novel tools by using the data reported in this thesis as discussed below:

5.2.1 Towards anti-phishing defense in Small and Medium Enterprises (SMEs)

Small and Medium Enterprises (SMEs) may have unique characteristics, operational structures, and resource constraints [168]. According to ACSC, "*Australian small to medium businesses (SMBs) operate in a different environment compared to larger enterprises, with 97% of Australian businesses having less than 20 staff*" [169]. Managing security threats and vulnerabilities effectively for SMEs is increasingly challenging [170].

Chapter 4 of this thesis reveals that SMEs often suffer from budget constraints in adopting appropriate phishing tools as mentioned by a practitioner in our study: "*well, the first major thing [while adopting an anti-phishing tool for the organization] that comes into it is budget which kind of sucks... So then we'll nail it down to well, let's just say 4 options. And then, after that, we'll do an extensive analysis of those. And then we can do like an options matrix and then sort of work out which one's best fit for us*" [P10]. Again, practitioner P9, who is the CEO of an SME, mentioned that they do not have any anti-phishing solutions deployed in their organizations. These data highlight that the requirements of SMEs are worth further investigation. Understanding their requirements would allow for the development of phishing defense strategies that are specifically tailored to these organizations' size, structure, and operational dynamics. Requirement analysis for anti-phishing defense can help identify the available resources, including budget, personnel, and technology required for a particular type of organization to enable the implementation of cost-effective and feasible defense measures. It can help assess the organization's risk profile, taking into account factors such as the nature of the business, industry regulations, and potential consequences of a phishing attack.

Organizations often rely on third-party vendors for anti-phishing solutions, the cost of which can amount to several thousand dollars [66], yet they could end up adopting anti-phishing solutions not suitable for them. For example, the study conducted by Brunken et. al [66] found that the adopted phishing simulation campaign from external vendors did not fit the cultural sensitivities and organizational policies of the organizations examined in that study. Consequently, a phishing simulation campaign would not be an effective intervention

unless the requirements of the organizations and their settings are well understood. In this regard, the outcome of requirement analysis can assist these organizations in selecting vendors and solutions that match their needs, budget constraints, and technical capabilities. Again, SMEs may experience growth and changes in their business landscape. Requirement analysis can ensure that phishing defense strategies are scalable and can adapt to the evolving needs of the organization.

5.2.2 Automated tool for presenting the guidelines to the practitioners

Features collected from the practitioners for a tool to access the guidelines (discussed in Chapter 4) indicate that it would be beneficial for the majority of the practitioners to design an automated tool using machine learning and natural language processing techniques that can generate the guidelines based on practitioners input requirements (an example study in other domain: [171]). In this way, practitioners can write their requirements in text format which can be processed to automatically generate a guideline relevant to them. This would be useful and can provide flexibility as practitioners do not need to be very familiar with the tool options to get a desired guideline.

To achieve this, first, a database of phishing intervention design, implementation, and evaluation guidelines for the practitioners can be created. Since our 41 guidelines might not be enough to train a machine learning model, other recommendations can be extracted through document analysis or by systematically collecting following a similar approach we have taken in our multivocal literature review discussed in Chapter 2 or data augmentation techniques can be used to improve the data availability constraints. This database of guidelines can be used to train a machine-learning model to generate guidelines that match practitioners' search queries. We demonstrated an example of a conceptual framework of the automated guideline generation system in Figure 5.1. Please note that the mapping (between guidelines and interventions) shown in Step 1 can be performed in terms of other aspects. For example, in addition to mapping guidelines with intervention types, mapping the intervention stages with the guidelines can be considered.

5.2.3 Architecture centric guidelines

Our study reported in Chapter 4 indicates that sometimes practitioners play multiple roles and work on different stages of phishing interventions. For example, one of the practitioners in our study mentioned: "*As was the case for me, practitioners manage multiple roles, especially in relation to security. For example, I had to perform tasks related to information security, risk management, policy setup, software security (related to our services), etc. This is mostly the case with smaller companies where one security engineer does most of the above tasks. Therefore, if a person selects a role, that person selects only one security-related responsibility in the organization to view the guidelines*" [P6].

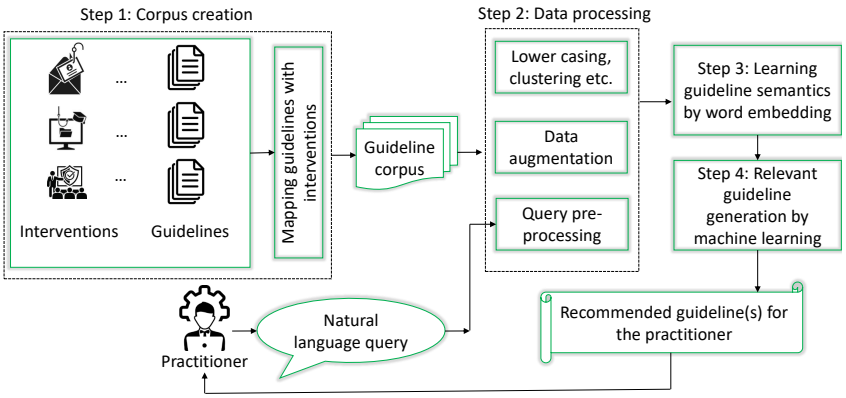

FIGURE 5.1. A conceptual framework for the automated guideline generation based on practitioners' input

Therefore, future studies can extend our role-based and stage-based personalization of the guidelines to produce customization in terms of other aspects. One approach would be to personalize the guidelines based on the architectural components of the phishing interventions. In the context of anti-phishing measures, architecture-centric guidelines refer to a structured approach that emphasizes the integration and alignment of the guidelines within the overall architecture of an anti-phishing system or tool adopted by the organization or the overall organization processes. Future researchers can first identify the functional and non-functional components of different types of interventions and then develop or categorize guidelines based on these architectural components.

Customizing the guidelines based on the architectural elements of diverse interventions can offer the following advantages: (1) discerning the guidelines based on the functional components facilitates the seamless integration of anti-phishing defense mechanisms into pre-existing cyber security systems and workflows, mitigating disruptions and fostering a cohesive security infrastructure; (2) it ensures that guidelines encompass a broad spectrum of aspects contributing to the development of a more comprehensive defense strategy.

5.2.4 Rigorous investigation of our reported challenges in large-scale settings

Chapter 2 and Chapter 4 revealed the challenges identified from the literature and in practice respectively. These challenges open the opportunity for future researchers to investigate most of these challenges in large-scale settings in detail. This is particularly important as it would create greater ecological validity of the challenges reported as well as confirm to what extent these challenges exist in real-world settings.

For example, our identified challenge, Ch15, highlighted the limitations of the current anti-phishing planning, policies, and guidelines in the organizations, including outdated recommendations in the organizations' website for the end-users, lack of formal procedures to invoke behavior changes, outdated phishing tools for IT management, etc (Section 2.5.2). A further investigation (interview

study, case study or survey) can be performed by discussing with the practitioners who are involved in dealing with these tasks. In this regard, some possible interview questions could be what procedures are in place in the organization to encourage employees to behave securely, how often the anti-phishing recommendations on the organizations' websites are updated, and what are the challenges behind updating these recommendations, the source of the presented recommendations, what are the driving factors organizations consider to select an anti-phishing tool and what is the obstacle to adopt them, etc. The outcome of these challenges can be communicated to the responsible authority, which can potentially improve the design, implementation, and evaluation of the phishing interventions. For instance, a study [43] identifying some issues in email services like Gmail has contacted them about their (Gmail) UI interface limitations. Later on, the authors found that the Gmail team had updated their interfaces, and some of the issues raised were resolved.

5.2.5 Evaluation of the updated guidelines

In Chapter 4, we reported the strengths and weaknesses of our guidelines from practitioners' perspectives. The identified weaknesses serve as a basis for potential updates to the guidelines, facilitating a subsequent evaluation of the revised guidelines through engagement with a distinct cohort of practitioners. The subsequent evaluation of the updated guidelines serves as a valuable step in the ongoing refinement process, ensuring that the recommendations evolve to meet the dynamic challenges faced by practitioners. It fosters a responsive and adaptive approach to guideline development, leading to more effective and user-centric outcomes. Conducting a subsequent evaluation can bring in several advantages:

- It would allow for an iterative process of improvement. It would provide an opportunity to refine and enhance the guidelines based on the insights gained from the initial evaluation.
- The evaluation can serve as a validation mechanism for the updates made to the guidelines. It may help assess whether the modifications effectively address the identified weaknesses and contribute to overall improvement.
- Engaging with a different set of practitioners would allow for the incorporation of diverse perspectives and feedback. This can contribute to a more comprehensive understanding of the guidelines' effectiveness and usability.
- Evaluating the guidelines with a different group of practitioners would help test the generalizability of the guidelines. It would help assess whether the guidelines are applicable across various practitioner backgrounds, contexts, and preferences.
- The subsequent evaluation may reveal new insights and challenges that were not apparent in the initial assessment. This expanded understanding can contribute to a more nuanced and comprehensive set of guidelines.
- Assessing the guidelines with a diverse set of practitioners would help gauge practitioners' acceptance and adoption. It can provide insights into how well the updated guidelines align with the practical needs and preferences of the practitioners.

- Through iterative evaluations, the guidelines have the potential to become more tailored and effective. Fine-tuning based on practitioner feedback can contribute to generating guidelines that are better aligned with real-world scenarios.

Appendix A

List of Selected Academic Studies

Table A.1. List of academic primary studies

No.	Title	Authors	Venue	Year	PETA types	Rank
P1	You've been warned: An empirical study of the effectiveness of web browser phishing warnings	Egelman S., Cranor L.F., Hong J.	Human Factors in Computing Systems	2008	Awareness	A*
P2	Do security toolbars actually prevent phishing attacks?	Wu M., Miller R.C., Garfinkel S.L.	Human Factors in Computing Systems	2006	Awareness	A*
P3	The emperor's new security indicators an evaluation of website authentication and the effect of role playing on usability studies	Schechter S.E., Dhamija R., Ozment A., Fischer I.	IEEE Symposium on Security and Privacy	2007	Awareness	A*
P4	Alice in warningland: A large-scale field study of browser security warning effectiveness	Akhawe D., Felt A.P.	USENIX Security Symposium	2013	Awareness	A*
P5	Protecting people from phishing: The design and evaluation of an embedded training email system	Kumaraguru P., Rhee Y., Acquisti A., Cranor L.F., Hong J., Nunge E.	Human Factors in Computing Systems	2007	Training	A*
P6	Security awareness of computer users: A phishing threat avoidance perspective	Arachchilage N.A.G., Love S.	Computers in Human Behavior	2014	Education	B
P7	Going spear phishing: Exploring embedded training, and awareness	Caputo D.D., Pfleeger S.L., Freeman J.D., Johnson M.E.	IEEE Security and Privacy Magazine	2014	Training	B
P8	Why phishing still works: User strategies for combating phishing attacks	Alsharnouby M., Alaca F., Chiasson S.	International Journal of Human Computer Studies	2015	Awareness	A
P9	A game design framework for avoiding phishing attacks	Arachchilage N.A.G., Love S.	Computers in Human Behavior	2013	Education	B
P10	Phishing threat avoidance behaviour: An empirical investigation	Arachchilage N.A.G., Love S., Beznosov K.	Computers in Human Behavior	2016	Education	B
P11	Security education against Phishing: A modest proposal for a Major Rethink	Kirlappos I., Sasse M.A.	IEEE Security and Privacy Magazine	2012	Education	B
P12	Phishing for phishing awareness	Jansson K., Von Solms R.	Behaviour and Information Technology	2013	Training	B
P13	Priming and warnings are not effective to prevent social engineering attacks	Junger M., Montoya L., Overink F.-J.	Computers in Human Behavior	2017	Awareness	B
P14	An experience sampling study of user reactions to browser warnings in the field	Reeder R., Felt A.P., Consolvo S., Malkin N., Thompson C., Egelman S.	Human Factors in Computing Systems	2018	Awareness	A*

Continued on next page

List of academic primary studies – continued from previous page

No.	Title	Authors	Venue	Year	PETA types	Rank
P15	Who provides phishing training? Facts, stories, and people like me	Wash R., Cooper M.M.	Human Factors in Computing Systems	2018	Training	A*
P16	End-to-end measurements of email spoofing attacks	Hu H., Wang G.	USENIX Security Symposium	2018	Awareness	A*
P17	A multi-modal neuro-physiological study of phishing detection and malware warnings	Neupane A., Rahman Md.L., Saxena N., Hirshfield L.	Computer and Communications Security	2015	Awareness	A*
P18	Towards a contingency approach with whitelist- and blacklist-based anti-phishing applications: What do usability tests indicate?	Li L., Berki E., Helenius M., Ovaska S.	Behaviour and Information Technology	2014	Awareness	B
P19	What.Hack: Engaging Anti-Phishing Training through a Role-playing Phishing Simulation Game	Wen Z.A., Lin Z., Chen R., Andersen E.	Human Factors in Computing Systems	2019	Training	A*
P20	Towards preventing QR code based attacks on android phone using security warnings	Yao H., Shin D.	ACM Asia Conference on Computer and Communications Security	2013	Awareness	A*
P21	How effective is anti-phishing training for children?	Lastdrager E., Gallardo I.C., Hartel P., Junger M.	Symposium on Usable Privacy and Security	2019	Training	B
P22	An Empirical Evaluation of Security Indicators in Mobile Web Browsers	Amrutkar C., Traynor P., Van Oorschot P.C.	IEEE Transactions on Mobile Computing	2015	Awareness	A*
P23	Building anti-phishing browser plug-ins: An experience report	Raffetseder T., Kirda E., Kruegel C.	International Workshop on Software Engineering for Secure Systems	2007	Awareness	A*
P24	Fishing for phishers. Improving Internet users' sensitivity to visual deception cues to prevent electronic fraud	Moreno-Fernández M.M., Blanco F., Garaizar P., Matute H.	Computers in Human Behavior	2017	Training	B
P25	Put your warning where your link is: Improving and evaluating email phishing warnings	Petelka J., Zou Y., Schaub F.	Human Factors in Computing Systems	2019	Awareness	A*
P26	Spear phishing in a barrel: Insights from a targeted phishing campaign	Burns A.J., Johnson M.E., Caputo D.D.	Journal of Organizational Computing and Electronic Commerce	2019	Training	B
P27	Evaluation of a mandatory phishing training program for high-risk employees at a US healthcare system	Gordon W.J., Wright A., Glynn R.J., Kadakia J., Mazzone C., Leinbach E., Landman A.	Journal of the American Medical Informatics Association	2019	Training	A
P28	Phishy - A serious game to train enterprise users on phishing awareness	Gokul C.J., Pandit S., Vaddepalli S., Tupsamudre H., Banahatti V., Lodha S.	Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts	2018	Training	A*
P29	The design and evaluation of a theory-based intervention to promote security behaviour against phishing	Jansen J., van Schaik P.	International Journal of Human Computer Studies	2019	Awareness	A
P30	Impact of security awareness training on phishing click-through rates	Carella A., Kotsoev M., Truta T.M.	IEEE International Conference on Big Data	2017	Training	B

Continued on next page

List of academic primary studies – continued from previous page

No.	Title	Authors	Venue	Year	PETA types	Rank
P31	Evaluation of personalized security indicators as an anti-phishing mechanism for smartphone applications	Marforio C., Masti R.J., Soriente C., Kostiainen K., Čapkun S.	Human Factors in Computing Systems	2016	Awareness	A*
P32	Measuring the effectiveness of embedded phishing exercises	Siadati H., Palka S., Siegel A., McCoy D.	USENIX Workshop on Cyber Security Experimentation and Test (co-located with USENIX Security symposium)	2017	Training	A*
P33	The description-experience gap in the effect of warning reliability on user trust and performance in a phishing-detection context	Chen J., Mishler S., Hu B., Li N., Proctor R.W.	International Journal of Human Computer Studies	2018	Awareness	A
P34	An investigation of phishing awareness and education over time: When and how to best remind users	Reinheimer B., Aldag L., Mayer P., Mossano M., Duezguen R., Lofthouse B., von Landesberger T., Volkamer M.	Symposium on Usable Privacy and Security	2020	Training	B
P35	Investigating Teenagers' Ability to Detect Phishing Messages	Nicholson J., Javed Y., Dixon M., Coventry L., Ajayi O.D., Anderson P.	IEEE European Symposium on Security and Privacy Workshops	2020	Education	A*
P36	Engaging users with educational games: The case of phishing	Dixon M., Nicholson J., Arachchilage N.A.G.	Human Factors in Computing Systems	2019	Education	A*
P37	How persuasive is a phishing email? A phishing game for phishing awareness	Fatima R., Yasin A., Liu L., Wang J.	Journal of Computer Security	2019	Education	B
P38	Employees' Behavior in Phishing Attacks: What Individual, Organizational, and Technological Factors Matter?	Shahbaznezhad H., Kolini F., Rashidirad M.	Journal of Computer Information Systems	2021	Training	B
P39	Intelligent explanation generation system for phishing webpages by employing an inference system	Ramesh G., Selvakumar K., Venugopal A.	Behaviour and Information Technology	2017	Awareness	B
P40	Deterrent effects of punishment and training on insider security threats: a field experiment on phishing attacks	Kim B., Lee D.-Y., Kim B.	Behaviour and Information Technology	2020	Training	B
P41	How Experts Detect Phishing Scam Emails	Wash R.	ACM Human-Computer Interaction (/CSCW)	2020	Education	A
P42	Analysis of publicly available anti-phishing webpages: Contradicting information, lack of concrete advice and very narrow attack vector	Mossano M., Vaniea K., Aldag L., Duzgun R., Mayer P., Volkamer M.	IEEE European Symposium on Security and Privacy Workshops, Euro S and PW 2020	2020	Education	A*
P43	Forcing Johnny to login safely	Herzberg A., Margulies R.	Journal of Computer Security	2013	Awareness	B
P44	Training johnny to authenticate (Safely)	Herzberg A., Margulies R.	IEEE Security and Privacy Magazine	2012	Awareness	B
P45	I don't need an expert! making url phishing features human comprehensible	Althobaiti K., Meng N., Vaniea K.	Human Factors in Computing Systems	2021	Education	A*

Continued on next page

List of academic primary studies – continued from previous page

No.	Title	Authors	Venue	Year	PETA types	Rank
P46	Be the Phisher - Understanding Users' Perception of Malicious Domains	Quinkert F., Degeling M., Blythe J., Holz T.	ACM Asia Conference on Computer and Communications Security	2020	Education	A*
P47	To click or not to click is the question: Fraudulent URL identification accuracy in a community sample	Pearson E., III, Bethel C.L., Jarosz A.F., Berman M.E.	IEEE International Conference on Systems, Man, and Cybernetics	2017	Education	B
P48	Facts and stories in phishing training: A replication and extension	Marsden J., Albrecht Z., Berggren P., Halbert J., Lemons K., Moncivais A., Thompson M.	Human Factors in Computing Systems	2020	Training	A*
P49	UI-dressing to detect phishing	Iacono L.L., Nguyen H.V., Hirsch T., Baiers M., Moller S.	IEEE International Conference on High Performance Computing and Communications	2014	Awareness	A
P50	A Case Study of Phishing Incident Response in an Educational Organization	Althobaiti K., Jenkins A.D.G., Vaniea K.	Proceedings of the ACM on Human-Computer Interaction (/CSCW)	2021	Education	A
P51	Knowledge and capabilities that non-expert users bring to phishing detection	Wash R., Nthala N., Rader E.	Symposium on Usable Privacy and Security	2021	Education	B
P52	Examining Factors Impacting the Effectiveness of Anti-Phishing Trainings	Sumner A., Yuan X., Anwar M., McBride M.	Journal of Computer Information Systems	2021	Training	B
P53	Simulated Phishing Attack and Embedded Training Campaign	Yeoh W., Huang H., Lee W.-S., Al Jafari F., Mansson R.	Journal of Computer Information Systems	2021	Training	B

Appendix B

List of Selected Grey Studies

TABLE B.1. List of grey primary studies

No.	Title	Author	Page	Link	Date	PETA type	Tier
P54	Phishing defense and govenance, how to improve user awareness, enhance controls and build process maturity	ISACA	1	https://terranovasecurity.com/white-papers/	2019	Training, awareness	1st
P55	How to avoid phishing simulations false positives?	Terranova-security	1	https://terranovasecurity.com/phishing-simulations-false-positives/	27-Jan	Training	3rd
P56	Why is phishing awareness training important?	Terranova-security	1	https://terranovasecurity.com/why-is-phishing-training-so-important/	20-Aug	Training	3rd
P57	Gone phishing tournament, phishing benchmark, global report	Terranova-security	website crawling	https://terranovasecurity.com/gpt-thank-you/	2021	Training	1st
P58	2022 state of the Phish, an in-depth exploration of user awareness,vulnerability and resilience	Proofpoint	1	https://www.proofpoint.com/us/resources/threat-reports/state-of-phish	2022	Training	1st
P59	How to transform employee worst, practices into enterprise best practices	KnowB4	1	https://info.knowbe4.com/whitepaper-employee-worst-best-practices-enterprise-security?hsLang=en	-	Training	1st
P60	Example security awareness, training policy guide	KnowB4	1	https://info.knowbe4.com/wp-example-sat-policy-guide?hsLang=en	-	Training	1st
P61	Building an effective and comprehensive, security awareness program	KnowB4	1	https://info.knowbe4.com/wp-building-effective-comprehensive-sat?hsLang=en	-	Training	1st
P62	Buyers guide to phishing training	Hoxhunt	4	https://www.hoxhunt.com/ebooks/the-buyers-guide-to-phishing-training	2020	Training	3rd
P63	How eckes-granini group transformed cybersecurity awareness with Hoxhunt	Hoxhunt	4	https://www.hoxhunt.com/case-studies/how-eckes-granini-group-transformed-cybersecurity-awareness-with-hoxhunt	-	Training	2nd

Continued on next page

List of grey primary studies – continued from previous page

No.	Title	Author	Page	Link	Date	PETA type	Tier
P64	Ramboll educates employees on email-based threats with Hoxhunt	Hoxhunt	4	https://www.hoxhunt.com/case-studies/ramboll-educates-employees-on-email-based-threats-with-hoxhunt	-	Training	2nd
P65	Ordermark built to last with cybersecurity awareness as foundation	Hoxhunt	4	https://www.hoxhunt.com/case-studies/case-study-ordermark-is-built-to-last-with-cybersecurity-awareness-as-a-cultural-foundation	-	Training	2nd
P66	How agricultural technology leader, Kverneland group sewed awareness training and reaped resilience	Hoxhunt	4	https://www.hoxhunt.com/case-studies/how-kverneland-group-sewed-awareness-training-and-reaped-resilience	-	Training	2nd
P67	5 Tips for evaluating phishing simulation solutions	PhishLabs	10	https://www.phishlabs.com/blog/5-tips-for-evaluating-phishing-simulation-solutions/	2/17/2016	Training	3rd
P68	How to run simulated phishing campaigns	Agari	16	https://www.agari.com/email-security-blog/how-to-run-simulated-phishing-campaigns/	1/5/2021	Training	3rd
P69	Best practice phishing simulation	SoSafe	16	https://sosafe-awareness.com/resources/guides/bpr-phishing/	2019	Training	1st

Appendix C

Data Extraction Form

Table C.1. Data extraction form used in this MLR

ID	Literature	Data type	Description	RQ
D1	Academic	Title	The title of the paper	Demographic data
D2	Academic	Author(s)	The author(s) of the paper	Demographic data
D3	Academic	Venue	The publication venue	Demographic data
D4	Academic	Year	The year of the publication	Demographic data
D5	Academic	Publication type	The type of the publication	Demographic data
D6	Academic	Rank	CORE ranking of the publication venue	Demographic data
D7	Grey	Title	The title of the source	Demographic data
D8	Grey	Author/ Organisation	Name of the author/organisation	Demographic data
D9	Grey	Date	The date of the source (if available)	Demographic data
D10	Grey	Link	The link of the source	Demographic data
D11	Grey	Outlet type	The tier type of the source	Demographic data
D12	Grey	Page no.	Google page number of the source	Demographic data
D13	Academic, Grey	Challenge(s)	The challenge reported in the article in the design, implementation, and evaluation stages of PETA	RQ1

Continued on next page

Data extraction form – continued from previous page

ID	Literature	Data type	Description	RQ
D14	Academic, Grey	Critical Factor(s)	Factor(s) discussed or Recommendation(s)/Guideline(s) provided in the article to be effective in the design, implementation, and evaluation stages to improve the success of PETA	RQ2
D15	Academic, Grey	Limitation(s)	Limitation discussed in the article	Discussion
D16	Academic	Future work	The reported future work in the study	Discussion

Appendix D

Demographics Collection Form

1. Please write your full name .....
2. Please provide your contact email .....
3. Please select your role in your organization
   O Chief information security officer
   O Member of the security team
   O Software developer
   O System designer
   O Web developer
   O UI/UX designer
   O Manager
   O Head of the organization
   O Others (please specify) .....
4. Please write the country name your organization is in .....
5. In which domain does your company operate?
   O Global Growth and Operations
   O Energy
   O Capital
   O Healthcare
   O Aviation
   O Transportation
   O Home and Business Solutions
   O Finance
   O Human Resources
   O Commercial, Public Relations
   O Legal
   O Business Development
   O Global Research
   O Defense
   O Others (Please specify) .....
6. What is the size of your organization (number of employees)?
   O 10 to 49
   O 50 to 249
   O 250 to 1000
   O More than 1000

7. How long have you been in the industry (total industry experience)?
   O Less than 1 year
   O 1 to 3 years
   O 3 to 5 years
   O 5 to 10 years
   O More than 10 years

8. Please select your highest level of academic qualification.
   O High school or less
   O Bachelor's degree
   O Master's degree
   O PhD degree
   O Others (Please specify) .....

9. Please describe your knowledge and experience in phishing [please select all that apply].

   O I have no experience dealing with or mitigating phishing attacks.
   O I have designed anti-phishing technologies or solutions.
   O I have been involved in the incident response process for phishing attacks.
   O I participate in creating or implementing phishing awareness programs within my organization
   O I have designed phishing awareness intervention
   O I have conducted user education sessions to improve phishing awareness
   O I have experience in overseeing and managing cyber security teams and strategies within an organization
   O Others (Please specify) .....

Appendix E

Interview Questions

E.1 Understanding the current anti-phishing practices

1. What anti-phishing techniques and mechanisms are employed by your organization to protect employees from phishing attacks?
   - What phishing education and training methods are in place to educate employees on phishing attacks?
   - How often does your organization perform phishing training?
   - Does your organization send notifications to the employees before conducting a phishing simulation campaign?
   - Does your organization take feedback from employees to update the anti-phishing policies and procedures?
   - How does your organization test the knowledge level of the employees after phishing training?
   - What actions are taken during real phishing attacks?

E.2 Understanding the current challenges

1. What challenges do you typically face in designing anti-phishing solutions (e.g., tools, interventions, or technology)? [For designers/developers]
   What challenges do you typically face when combating phishing attacks in your organization? [For security team members/C-suite employees/Cybersecurity experts]
   - Please share some examples of such challenges you have experienced.
   - Why do you think such challenges that you mentioned are challenging?

E.3 Evaluation of the guidelines

1. What do you think are the strengths of the guidelines?

2. Please rate the following statements regarding the strengths of the guidelines.
   "The guidelines are reliable as they are systematically organized from trusted source":

O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The guidelines convey meaningful and important information":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The guidelines are useful because I, as a practitioner, can find information in a reasonable timeframe":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The guidelines can help prioritize and resolve such human-centric issues in phishing interventions more effectively":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The guidelines are clearly stated and easily understandable":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The guidelines can help address the current and relevant phishing threats":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The guidelines can improve the outcome of anti-phishing interventions by reducing their usability issues":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

3. Besides the options provided, do you have any other comments regarding the strengths of the guidelines?
   .....

4. Did you encounter any difficulties or confusion while trying to follow the guidelines presented to you by the tool? [Please select all that apply]
   O I have difficulty understanding the language and terminology used in the guidelines
   O I find the guidelines are challenging to interpret or comprehend
   O I find the guidelines unclear and difficult to follow
   O I need additional explanatory notes or examples to follow the guidelines
   O I have used similar guidelines before which was easier to follow and helpful
   O I have not encountered any difficulty or confusion in following any of the guidelines
   O Others (Please specify) .....

5. Were there any constraints or challenges that you anticipate in implementing the guidelines within your organization?
   O The guidelines lacked practical guidance to implement within my organizations' resources and capabilities
   O It is difficult to adapt the guidelines to suit the unique requirements of my organization
   O It is difficult to integrate the recommendations with the existing security measures or protocols in my organization
   O The guidelines did not provide actionable steps to implement
   O Others (Please specify) .....

6. What is your recommendation to improve the guidelines to overcome current phishing threats?

E.4 Evaluation of PhishGuide and collecting desired features for an envisioned tool

1. Provide your opinion about the best aspects of the tool.

2. Please rate the following statements regarding the strengths of the tool.
   "I can get guideline(s) in a straightforward manner using this tool":
   O Strongly disagree (1)
   O Disagree (2)
   O Neutral (3)
   O Agree (4)
   O Strongly agree (5)

   "The organization of the options seems quite logical":
   O Strongly disagree (1)
   O Disagree (2)

O Neutral (3)
O Agree (4)
O Strongly agree (5)

"There are a reasonable number of steps required to get a guideline":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The tool presents itself in a very attractive way":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The way that system information is presented is clear and understandable":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"It is easy to see at a glance what the options are at each stage":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"I enjoyed my time using the tool as the tool could provide guidelines personalized to me.":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The tool has made the guidelines on anti-phishing interventions easily accessible to me":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The tool has saved my time required for searching and accessing guidelines that are scattered throughout various online sources":

O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

Others (Please specify) .....

3. Provide your opinion about the challenges or difficulties you faced while using the tool.

"The tool hasn't always done what I was expecting":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"There is too much to read before I can use the tool":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"It takes too long to learn the tool functions":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"I keep having to go back to look at the help information":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The tool did not provide definitions of the terms used":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"It is easy to forget how to do things with this tool":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"Learning how to use the tool is difficult":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"I have to look for assistance most times when I use this tool":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"There is not enough information on the tool":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

"The help information given by this tool was not helpful":
O Strongly disagree (1)
O Disagree (2)
O Neutral (3)
O Agree (4)
O Strongly agree (5)

Others (Please specify) .....

4. What are your recommendations to improve the tool to meet your needs, preferences, and expectations?

5. In your opinion, what features an anti-phishing guideline generator tool like ours should consist of to facilitate the enhanced design, implementation, and evaluation of anti-phishing interventions?

Appendix F

Ethics Approval Form

The ethics approval form from The University of Adelaide for the interview study is displayed on the next page.

RESEARCH SERVICES
OFFICE OF RESEARCH ETHICS, COMPLIANCE AND INTEGRITY
THE UNIVERSITY OF ADELAIDE

LEVEL 3, RUNDLE MALL PLAZA
50 RUNDLE MALL
ADELAIDE SA 5000 AUSTRALIA

TELEPHONE +61 8 8313 5137
FACSIMILE +61 8 8313 3700
EMAIL hrec@adelaide.edu.au

CRICOS Provider Number 00123M

Our reference 37109

20 June 2023

Dr Asangi Jayatilaka
Computer Science

Dear Dr Jayatilaka

ETHICS APPROVAL No: H-2023-124
PROJECT TITLE: Evaluation of personalized guidelines for the design, implementation, and evaluation of anti-phishing interventions

The ethics application for the above project has been reviewed by the Executive, Human Research Ethics Committee and is deemed to meet the requirements of the *National Statement on Ethical Conduct in Human Research 2007 (Updated 2018)* involving no more than low risk for research participants.

You are authorised to commence your research on: 20/06/2023
The ethics expiry date for this project is: 30/06/2026

NAMED INVESTIGATORS:

Chief Investigator:	Dr Asangi Jayatilaka
Student - Postgraduate Doctorate by Research (PhD):	Miss Orvila Sarker
Associate Investigator:	Associate Professor Chelsea Liu
Associate Investigator:	Dr Sherif Haggag
Associate Investigator:	Professor Ali Babar

CONDITIONS OF APPROVAL: Thank you for your considered responses to the matters raised. The revised application provided on 16.06.2023 has been approved.

Ethics approval is granted for three years and is subject to satisfactory annual reporting. The form titled Annual Report on Project Status is to be used when reporting annual progress and project completion and can be downloaded at http://www.adelaide.edu.au /research-services/oreci/human/reporting/. Prior to expiry, ethics approval may be extended for a further period.

Participants in the study are to be given a copy of the information sheet and the signed consent form to retain. It is also a condition of approval that you immediately report anything which might warrant review of ethical approval including:

- serious or unexpected adverse effects on participants,
- previously unforeseen events which might affect continued ethical acceptability of the project,
- proposed changes to the protocol or project investigators; and
- the project is discontinued before the expected date of completion.

Yours sincerely,

Dr Tiffany Gill
Acting Chair

The University of Adelaide

Bibliography

[1] P. Langlois, "2020 data breach investigations report," 2020.

[2] C. Crowley and J. Pescatore, "Common and best practices for security operations centers: Results of the 2019 soc survey," *SANS, Bethesda, MD, USA, Tech. Rep*, 2019.

[3] Proofpoint. "2023 state of phish." (2022), [Online]. Available: https://www.proofpoint.com/sites/default/files/threat-reports/pfpt-us-tr-state-of-the-phish-2023.pdf.

[4] APWG, *Phishing activity trends report*, 2022. [Online]. Available: https://apwg.org/trendsreports/.

[5] M. M. Alani and H. Tawfik, "Phishnot: A cloud-based machine-learning approach to phishing url detection," *Computer Networks*, p. 109 407, 2022.

[6] S. Magdy, Y. Abouelseoud, and M. Mikhail, "Efficient spam and phishing emails filtering based on deep learning," *Computer Networks*, vol. 206, p. 108 826, 2022.

[7] D.-J. Liu, G.-G. Geng, and X.-C. Zhang, "Multi-scale semantic deep fusion models for phishing website detection," *Expert Systems with Applications*, vol. 209, p. 118 305, 2022.

[8] S. E. Schechter, R. Dhamija, A. Ozment, and I. Fischer, "The emperor's new security indicators," in *2007 IEEE Symposium on Security and Privacy (SP'07)*, IEEE, 2007, pp. 51–65.

[9] N. A. G. Arachchilage, S. Love, and K. Beznosov, "Phishing threat avoidance behaviour: An empirical investigation," *Computers in Human Behavior*, vol. 60, pp. 185–197, 2016.

[10] J. Petelka, Y. Zou, and F. Schaub, "Put your warning where your link is: Improving and evaluating email phishing warnings," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–15.

[11] J. Chen, S. Mishler, B. Hu, N. Li, and R. W. Proctor, "The description-experience gap in the effect of warning reliability on user trust and performance in a phishing-detection context," *International Journal of Human-Computer Studies*, vol. 119, pp. 35–47, 2018.

[12] A. C. Tally, J. Abbott, A. M. Bochner, S. Das, and C. Nippert-Eng, "Tips, tricks, and training: Supporting anti-phishing awareness among mid-career office workers based on employees' current practices," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–13.

[13] M. Zwilling, G. Klien, D. Lesjak, Ł. Wiechetek, F. Cetin, and H. N. Basim, "Cyber security awareness, knowledge and behavior: A comparative study," *Journal of Computer Information Systems*, vol. 62, no. 1, pp. 82–97, 2022.

[14] J. Abawajy, "User preference of cyber security awareness delivery methods," *Behaviour & Information Technology*, vol. 33, no. 3, pp. 237–248, 2014.

[15] A. Aleroud and L. Zhou, "Phishing environments, techniques, and countermeasures: A survey," *Computers & Security*, vol. 68, pp. 160–196, 2017.

[16] G. Desolda, L. S. Ferro, A. Marrella, T. Catarci, and M. F. Costabile, "Human factors in phishing attacks: A systematic literature review," *ACM Computing Surveys (CSUR)*, vol. 54, no. 8, pp. 1–35, 2021.

[17] S. Hu, C. Hsu, and Z. Zhou, "Security education, training, and awareness programs: Literature review," *Journal of Computer Information Systems*, pp. 1–13, 2021.

[18] N. A. G. Arachchilage and S. Love, "A game design framework for avoiding phishing attacks," *Computers in Human Behavior*, vol. 29, no. 3, pp. 706–714, 2013.

[19] P. Kumaraguru, Y. Rhee, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "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*, 2007, pp. 905–914.

[20] D. Akhawe and A. P. Felt, "Alice in warningland: A {large-scale} field study of browser security warning effectiveness," in *22nd USENIX Security Symposium (USENIX Security 13)*, 2013, pp. 257–272.

[21] A. Franz, V. Zimmermann, G. Albrecht, K. Hartwig, C. Reuter, A. Benlian, and J. Vogt, "{Sok}: Still plenty of phish in the sea—a taxonomy of {user-oriented} phishing interventions and avenues for future research," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, 2021, pp. 339–358.

[22] D. Lain, K. Kostiainen, and S. Čapkun, "Phishing in organizations: Findings from a large-scale and long-term study," in *2022 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2022, pp. 842–859.

[23] G. CJ, S. Pandit, S. Vaddepalli, H. Tupsamudre, V. Banahatti, and S. Lodha, "Phishy-a serious game to train enterprise users on phishing awareness," in *Proceedings of the 2018 annual symposium on computer-human interaction in play companion extended abstracts*, 2018, pp. 169–181.

[24] S. Sjouwerman, *Context is the key to phishing success*, 2021. [Online]. Available: https://blog.knowbe4.com/context-is-the-key-to-phishing-success.

[25] S. Sheng, B. Magnien, P. Kumaraguru, A. Acquisti, L. F. Cranor, J. Hong, and E. Nunge, "Anti-phishing phil: The design and evaluation of a game that teaches people not to fall for phish," in *Proceedings of the 3rd symposium on Usable privacy and security*, 2007, pp. 88–99.

[26] A. Caballero, "Security education, training, and awareness," in *Computer and information security handbook*, Elsevier, 2017, pp. 497–505.

[27] A. Jayatilaka, N. Beu, I. Baetu, M. Zahedi, M. A. Babar, L. Hartley, and W. Lewinsmith, "Evaluation of security training and awareness programs: Review of current practices and guideline," *arXiv preprint arXiv:2112.06356*, 2021.

[28] M. Bada, A. M. Sasse, and J. R. Nurse, "Cyber security awareness campaigns: Why do they fail to change behaviour?" *arXiv preprint arXiv:1901.02672*, 2019.

[29] B. Kaiser, J. Wei, E. Lucherini, K. Lee, J. N. Matias, and J. Mayer, "Adapting security warnings to counter online disinformation," in *30th USENIX Security Symposium (USENIX Security 21)*, 2021, pp. 1163–1180.

[30] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham, "School of phish: A real-world evaluation of anti-phishing training," in *Proceedings of the 5th Symposium on Usable Privacy and Security*, 2009, pp. 1–12.

[31] C. Reuter, L. L. Iacono, and A. Benlian, *A quarter century of usable security and privacy research: Transparency, tailorability, and the road ahead*, 2022.

[32] F. B. Salamah, M. A. Palomino, M. Papadaki, and S. Furnell, "The importance of the job role in social media cybersecurity training," in *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, 2022, pp. 454–462.

[33] J.-W. Bullee and M. Junger, "How effective are social engineering interventions? a meta-analysis," *Information & Computer Security*, 2020.

[34] C. Marforio, R. Jayaram Masti, C. Soriente, K. Kostiainen, and S. Čapkun, "Evaluation of personalized security indicators as an anti-phishing mechanism for smartphone applications," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 540–551.

[35] S. Egelman, L. F. Cranor, and J. Hong, "You've been warned: An empirical study of the effectiveness of web browser phishing warnings," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2008, pp. 1065–1074.

[36] M. Wu, R. C. Miller, and S. L. Garfinkel, "Do security toolbars actually prevent phishing attacks?" In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2006, pp. 601–610.

[37] M. Alsharnouby, F. Alaca, and S. Chiasson, "Why phishing still works: User strategies for combating phishing attacks," *International Journal of Human-Computer Studies*, vol. 82, pp. 69–82, 2015.

[38] R. Hoda, "Socio-technical grounded theory for software engineering," *IEEE Transactions on Software Engineering*, vol. 48, no. 10, pp. 3808–3832, 2021.

[39] R. Wash and M. M. Cooper, "Who provides phishing training? facts, stories, and people like me," in *Proceedings of the 2018 chi conference on human factors in computing systems*, 2018, pp. 1–12.

[40] J. Marsden, Z. Albrecht, P. Berggren, J. Halbert, K. Lemons, A. Moncivais, and M. Thompson, "Facts and stories in phishing training: A replication and extension," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–6.

[41] M. Dixon, N. A. Gamagedara Arachchilage, and J. Nicholson, "Engaging users with educational games: The case of phishing," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–6.

[42] L. Li, E. Berki, M. Helenius, and S. Ovaska, "Towards a contingency approach with whitelist-and blacklist-based anti-phishing applications: What do usability tests indicate?" *Behaviour & Information Technology*, vol. 33, no. 11, pp. 1136–1147, 2014.

[43] H. Hu and G. Wang, "End-to-end measurements of email spoofing attacks.," in *USENIX Security Symposium*, 2018, pp. 1095–1112.

[44] A. Sumner, X. Yuan, M. Anwar, and M. McBride, "Examining factors impacting the effectiveness of anti-phishing trainings," *Journal of Computer Information Systems*, vol. 62, no. 5, pp. 975–997, 2022.

[45] A. Herzberg and R. Margulies, "Training johnny to authenticate (safely)," *IEEE Security & Privacy*, vol. 10, no. 1, pp. 37–45, 2011.

[46] M. Mossano, K. Vaniea, L. Aldag, R. Düzgün, P. Mayer, and M. Volkamer, "Analysis of publicly available anti-phishing webpages: Contradicting information, lack of concrete advice and very narrow attack vector," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, 2020, pp. 130–139.

[47] K. Althobaiti, A. D. Jenkins, and K. Vaniea, "A case study of phishing incident response in an educational organization," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–32, 2021.

[48] D. D. Caputo, S. L. Pfleeger, J. D. Freeman, and M. E. Johnson, "Going spear phishing: Exploring embedded training and awareness," *IEEE Security & Privacy*, vol. 12, no. 1, pp. 28–38, 2013.

[49] A. Burns, M. E. Johnson, and D. D. Caputo, "Spear phishing in a barrel: Insights from a targeted phishing campaign," *Journal of Organizational Computing and Electronic Commerce*, vol. 29, no. 1, pp. 24–39, 2019.

[50] Z. A. Wen, Z. Lin, R. Chen, and E. Andersen, "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*, 2019, pp. 1–12.

[51] K. Althobaiti, N. Meng, and K. Vaniea, "I don't need an expert! making url phishing features human comprehensible," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–17.

[52] C. Amrutkar, P. Traynor, and P. C. Van Oorschot, "An empirical evaluation of security indicators in mobile web browsers," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 889–903, 2013.

[53] R. W. Reeder, A. P. Felt, S. Consolvo, N. Malkin, C. Thompson, and S. Egelman, "An experience sampling study of user reactions to browser warnings in the field," in *Proceedings of the 2018 CHI conference on human factors in computing systems*, 2018, pp. 1–13.

[54] A Reeves, P Delfabbro, and D Calic, "Encouraging employee engagement with cybersecurity: How to tackle cyber fatigue," *SAGE open*, vol. 11, no. 1, p. 21 582 440 211 000 049, 2021.

[55] L. L. Iacono, H. V. Nguyen, T. Hirsch, M. Baiers, and S. Möller, "Uidressing to detect phishing," in *2014 IEEE Intl Conf on High Performance Computing and Communications, 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC, CSS, ICESS)*, IEEE, 2014, pp. 747–754.

[56] L. Bauer, C. Bravo-Lillo, L. F. Cranor, and E. Fragkaki, "Warning design guidelines (cmu-cylab-13-002)," 2013.

[57] L. A. Shepherd and K. Renaud, "How to design browser security and privacy alerts," *arXiv preprint arXiv:1806.05426*, 2018.

[58] M. Gáliková, V. Švábenskỳ, and J. Vykopal, "Toward guidelines for designing cybersecurity serious games," in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, 2021, pp. 1275–1275.

[59] N. Ebert, T. Schaltegger, B. Ambuehl, L. Schöni, V. Zimmermann, and M. Knieps, "Learning from safety science: A way forward for studying cybersecurity incidents in organizations," *Computers & Security*, p. 103 435, 2023.

[60] L. Colusso, C. L. Bennett, G. Hsieh, and S. A. Munson, "Translational resources: Reducing the gap between academic research and hci practice," in *Proceedings of the 2017 conference on designing interactive systems*, 2017, pp. 957–968.

[61] C. Sas, S. Whittaker, S. Dow, J. Forlizzi, and J. Zimmerman, "Generating implications for design through design research," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2014, pp. 1971–1980.

[62] E. Buie, C. J. Hooper, and A. Houssian, "Practice interaction: Building bridges, closing the gap," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 2013, pp. 2493–2496.

[63] D. Hillman, Y. Harel, and E. Toch, "Evaluating organizational phishing awareness training on an enterprise scale," *Computers & Security*, p. 103 364, 2023.

[64] M. De Bona and F. Paci, "A real world study on employees' susceptibility to phishing attacks," in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, 2020, pp. 1–10.

[65] T. Raffetseder, E. Kirda, and C. Kruegel, "Building anti-phishing browser plug-ins: An experience report," in *Third International Workshop on Software Engineering for Secure Systems (SESS'07: ICSE Workshops 2007)*, IEEE, 2007, pp. 6–6.

[66] L. Brunken, A. Buckmann, J. Hielscher, and M. A. Sasse, "To do this properly, you need more resources: The hidden costs of introducing simulated phishing campaigns," in *32nd USENIX Security Symposium (USENIX Security 23)*, 2023, pp. 4105–4122.

[67] P. Kumaraguru, S. Sheng, A. Acquisti, L. F. Cranor, and J. Hong, "Lessons from a real world evaluation of anti-phishing training," in *2008 eCrime Researchers Summit*, IEEE, 2008, pp. 1–12.

[68] P. Kumaraguru, Y. Rhee, S. Sheng, S. Hasan, A. Acquisti, L. F. Cranor, and J. Hong, "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*, 2007, pp. 70–81.

[69] A. Carella, M. Kotsoev, and T. M. Truta, "Impact of security awareness training on phishing click-through rates," in *2017 IEEE International Conference on Big Data (Big Data)*, IEEE, 2017, pp. 4458–4466.

[70] E. Lastdrager, I. C. Gallardo, P. Hartel, and M. Junger, "How effective is {anti-phishing} training for children?" In *Thirteenth symposium on usable privacy and security (soups 2017)*, 2017, pp. 229–239.

[71] O. Sarker, A. Jayatilaka, S. Haggag, C. Liu, and M. A. Babar, "A multivocal literature review on challenges and critical success factors of phishing education, training and awareness," *Journal of Systems and Software*, vol. 208, p. 111 899, 2024.

[72] K. K. Greene, M. Steves, M. F. Theofanos, J. Kostick, *et al.*, "User context: An explanatory variable in phishing susceptibility," in *in Proc. 2018 Workshop Usable Security*, 2018.

[73] E. Tom, A. Aurum, and R. Vidgen, "An exploration of technical debt," *Journal of Systems and Software*, vol. 86, no. 6, pp. 1498–1516, 2013.

[74] K. M. Benzies, S. Premji, K. A. Hayden, and K. Serrett, "State-of-the-evidence reviews: Advantages and challenges of including grey literature," *Worldviews on Evidence-Based Nursing*, vol. 3, no. 2, pp. 55–61, 2006.

[75] Q. Mahood, D. Van Eerd, and E. Irvin, "Searching for grey literature for systematic reviews: Challenges and benefits," *Research synthesis methods*, vol. 5, no. 3, pp. 221–234, 2014.

[76] T. Security, *How to avoid phishing simulations false positives?* 2022. [Online]. Available: https://terranovasecurity.com/phishing-simulations-false-positives/.

[77] D. Jampen, G. Gür, T. Sutter, and B. Tellenbach, "Don't click: Towards an effective anti-phishing training. a comparative literature review," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, pp. 1–41, 2020.

[78] R. Heartfield and G. Loukas, "A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks," *ACM Computing Surveys (CSUR)*, vol. 48, no. 3, pp. 1–39, 2015.

[79] P. Burda, T. Chotza, L. Allodi, and N. Zannone, "Testing the effectiveness of tailored phishing techniques in industry and academia: A field experiment," in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, 2020, pp. 1–10.

[80] S. Baki and R. Verma, "Sixteen years of phishing user studies: What have we learned?" *arXiv preprint arXiv:2109.04661*, 2021.

[81] S. Das, A. Kim, Z. Tingle, and C. Nippert-Eng, "All about phishing: Exploring user research through a systematic literature review," *arXiv preprint arXiv:1908.05897*, 2019.

[82] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," *Technical report, EBSE Technical Report EBSE-2007-01*, 2007.

[83] V. Garousi, M. Felderer, and M. V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," *Information and Software Technology*, vol. 106, pp. 101–121, 2019.

[84] M. Shahin, M. A. Babar, and L. Zhu, "Continuous integration, delivery and deployment: A systematic review on approaches, tools, challenges and practices," *IEEE Access*, vol. 5, pp. 3909–3943, 2017.

[85] B. Kitchenham, R. Pretorius, D. Budgen, O. P. Brereton, M. Turner, M. Niazi, and S. Linkman, "Systematic literature reviews in software engineering–a tertiary study," *Information and software technology*, vol. 52, no. 8, pp. 792–805, 2010.

[86] M. Zahedi, M. Shahin, and M. A. Babar, "A systematic review of knowledge sharing challenges and practices in global software development," *International Journal of Information Management*, vol. 36, no. 6, pp. 995–1019, 2016.

[87] M. Shahin, M. A. Babar, and M. A. Chauhan, "Architectural design space for modelling and simulation as a service: A review," *Journal of Systems and Software*, vol. 170, p. 110 752, 2020.

[88] V. Garousi, M. Felderer, and T. Hacaloğlu, "Software test maturity assessment and test process improvement: A multivocal literature review," *Information and Software Technology*, vol. 85, pp. 16–42, 2017.

[89] V. Garousi and M. V. Mäntylä, "When and what to automate in software testing? a multi-vocal literature review," *Information and Software Technology*, vol. 76, pp. 92–117, 2016.

[90] C. Islam, M. A. Babar, and S. Nepal, "A multi-vocal review of security orchestration," *ACM Computing Surveys (CSUR)*, vol. 52, no. 2, pp. 1–45, 2019.

[91] B.-J. Butijn, D. A. Tamburri, and W.-J. v. d. Heuvel, "Blockchains: A systematic multivocal literature review," *ACM Computing Surveys (CSUR)*, vol. 53, no. 3, pp. 1–37, 2020.

[92] G. Cascavilla, D. A. Tamburri, and W.-J. Van Den Heuvel, "Cybercrime threat intelligence: A systematic multi-vocal literature review," *Computers & Security*, vol. 105, p. 102 258, 2021.

[93] G. Sharma, *Fight the phish—see how microsoft learn can help*, https://techcommunity.microsoft.com/t5/microsoft-learn-blog/fight-the-phish-see-how-microsoft-learn-can-help/ba-p/2824122, 2021.

[94] CORE, *Core conference rankings 2021: Process followed and data considered*, https://drive.google.com/file/d/1bKa40nheaQ3zfuXu3jSpKIw5TnhK9USR/view, 2021.

[95] CORE, *Computer science conference rankings descriptions*, https://drive.google.com/file/d/1q21YeVIEDYyjKJ9WBPXTgbRH_reCnV12/view, 2021.

[96] E. Souza, A. Moreira, and M. Goulão, "Deriving architectural models from requirements specifications: A systematic mapping study," *Information and software technology*, vol. 109, pp. 26–39, 2019.

[97] R. Croft, Y. Xie, and M. A. Babar, "Data preparation for software vulnerability prediction: A systematic literature review," *IEEE Transactions on Software Engineering*, 2022.

[98] B. Sabir, F. Ullah, M. A. Babar, and R. Gaire, "Machine learning for detecting data exfiltration: A review," *ACM Computing Surveys (CSUR)*, vol. 54, no. 3, pp. 1–47, 2021.

[99] B. Kitchenham, "Procedures for performing systematic reviews," *Keele, UK, Keele University*, vol. 33, no. 2004, pp. 1–26, 2004.

[100] F. Q. Da Silva, A. L. Santos, S. Soares, A. C. C. França, C. V. Monteiro, and F. F. Maciel, "Six years of systematic literature reviews in software engineering: An updated tertiary study," *Information and Software Technology*, vol. 53, no. 9, pp. 899–913, 2011.

[101] A. Soneji, F. B. Kokulu, C. Rubio-Medrano, T. Bao, R. Wang, Y. Shoshitaishvili, and A. Doupé, ""flawed, but like democracy we don't have a better system": The experts' insights on the peer review process of evaluating security papers," in *2022 IEEE Symposium on Security and Privacy (SP)*, IEEE, 2022, pp. 1845–1862.

[102] N. Dissanayake, A. Jayatilaka, M. Zahedi, and M. A. Babar, "Software security patch management-a systematic literature review of challenges, approaches, tools and practices," *Information and Software Technology*, vol. 144, p. 106 771, 2022.

[103] L. Jiang, A. Jayatilaka, M. Nasim, M. Grobler, M. Zahedi, and M. A. Babar, "Systematic literature review on cyber situational awareness visualizations," *arXiv preprint arXiv:2112.10354*, 2021.

[104] R. J. Adams, P. Smart, and A. S. Huff, "Shades of grey: Guidelines for working with the grey literature in systematic reviews for management and organizational studies," *International Journal of Management Reviews*, vol. 19, no. 4, pp. 432–454, 2017.

[105] V. Garousi and M. Felderer, "Experience-based guidelines for effective and efficient data extraction in systematic reviews in software engineering," in *Proceedings of the 21st International Conference on Evaluation and Assessment in Software Engineering*, 2017, pp. 170–179.

[106] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[107] A. Sbaraini, S. M. Carter, R. W. Evans, and A. Blinkhorn, "How to do a grounded theory study: A worked example of a study of dental practices," *BMC medical research methodology*, vol. 11, no. 1, pp. 1–10, 2011.

[108] S. Scataglini and G. Paul, *DHM and Posturography*. Academic Press, 2019.

[109] N. C. R. L. Y. Teraguchi and J. C. Mitchell, "Client-side defense against web-based identity theft," *Computer Science Department, Stanford University. Available: http://crypto. stanford. edu/SpoofGuard/webspoof. pdf*, 2004.

[110] A. Herzberg and A. Gbara, "Trustbar: Protecting (even naive) web users from spoofing and phishing attacks," Cryptology ePrint Archive, Report 2004/155. http://eprint. iacr. org/2004/155, Tech. Rep., 2004.

[111] J. A. Chaudhry, S. A. Chaudhry, and R. G. Rittenhouse, "Phishing attacks and defenses," *International journal of security and its applications*, vol. 10, no. 1, pp. 247–256, 2016.

[112] T. Dybå and T. Dingsøyr, "Empirical studies of agile software development: A systematic review," *Information and software technology*, vol. 50, no. 9-10, pp. 833–859, 2008.

[113] A. Ampatzoglou, S. Bibi, P. Avgeriou, M. Verbeek, and A. Chatzigeorgiou, "Identifying, categorizing and mitigating threats to validity in software engineering secondary studies," *Information and Software Technology*, vol. 106, pp. 201–230, 2019.

[114] P. T. Metaxas, "Web spam, social propaganda and the evolution of search engine rankings," in *International Conference on Web Information Systems and Technologies*, Springer, 2009, pp. 170–182.

[115] S. W. Schuetz, Z. R. Steelman, and R. A. Syler, "It's not just about accuracy: An investigation of the human factors in users' reliance on anti-phishing tools," *Decision Support Systems*, p. 113 846, 2022.

[116] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," *arXiv preprint arXiv:1702.08608*, 2017.

[117] M. T. Dzindolet, S. A. Peterson, R. A. Pomranky, L. G. Pierce, and H. P. Beck, "The role of trust in automation reliance," *International journal of human-computer studies*, vol. 58, no. 6, pp. 697–718, 2003.

[118] G. Vilone and L. Longo, "Explainable artificial intelligence: A systematic review," *arXiv preprint arXiv:2006.00093*, 2020.

[119] O. Sarker, S. Haggag, A. Jayatilaka, and C. Liu, "Personalized guidelines for design, implementation, and evaluation of anti-phishing interventions," in *Proceedings of the 17th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2023, pp. 1–12.

[120] P. Jaferian, D. Botta, F. Raja, K. Hawkey, and K. Beznosov, "Guidelines for designing it security management tools," in *Proceedings of the 2nd ACM Symposium on Computer Human interaction For Management of information Technology*, 2008, pp. 1–10.

[121] S. Chiasson, P. van Oorschot, and R. Biddle, "Even experts deserve usable security: Design guidelines for security management systems," in *SOUPS Workshop on Usable IT Security Management (USM)*, 2007, pp. 1–4.

[122] *Scopus*, https://www.scopus.com/.

[123] *Core*, https://www.core.edu.au/.

[124] CORE, *Core conference rankings 2021: Process followed and data considered*, https://drive.google.com/file/d/1bKa40nheaQ3zfuXu3jSpKIw5TnhK9USR/view.

[125] K.-P. Yee and K. Sitaker, "Passpet: Convenient password management and phishing protection," in *Proceedings of the second symposium on Usable privacy and security*, 2006, pp. 32–43.

[126] U. Bhatt, A. Xiang, S. Sharma, A. Weller, A. Taly, Y. Jia, J. Ghosh, R. Puri, J. M. Moura, and P. Eckersley, "Explainable machine learning in deployment," in *Proceedings of the 2020 conference on fairness, accountability, and transparency*, 2020, pp. 648–657.

[127] S. R. Hong, J. Hullman, and E. Bertini, "Human factors in model interpretability: Industry practices, challenges, and needs," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW1, pp. 1–26, 2020.

[128] R. Tomsett, D. Braines, D. Harborne, A. Preece, and S. Chakraborty, "Interpretable to whom? a role-based model for analyzing interpretable machine learning systems," *arXiv preprint arXiv:1806.07552*, 2018.

[129] G. Dupont, "The dirty dozen errors in maintenance," in *The 11th symposium on human factors in maintenance and inspection: Human error in aviation maintenance*, 1997.

[130] D. Hidellaarachchi, J. Grundy, R. Hoda, and I. Mueller, "The influence of human aspects on requirements engineering-related activities: Software practitioners' perspective," *ACM Transactions on Software Engineering and Methodology*, 2022.

[131] J. K. Nwankpa and P. M. Datta, "Remote vigilance: The roles of cyber awareness and cybersecurity policies among remote workers," *Computers & Security*, vol. 130, p. 103 266, 2023.

[132] K. M. Alnifie and C. Kim, "Appraising the manifestation of optimism bias and its impact on human perception of cyber security: A meta analysis," *Journal of Information Security*, vol. 14, no. 2, pp. 93–110, 2023.

[133] F. Kamei, G. Pinto, I. Wiese, M. Ribeiro, and S. Soares, "What evidence we would miss if we do not use grey literature?" In *Proceedings of the 15th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2021, pp. 1–11.

[134] S. Hermawati and G. Lawson, "Establishing usability heuristics for heuristics evaluation in a specific domain: Is there a consensus?" *Applied ergonomics*, vol. 56, pp. 34–51, 2016.

[135] R. N. Rajapakse, M. Zahedi, and M. A. Babar, "An empirical analysis of practitioners' perspectives on security tool integration into devops," in *Proceedings of the 15th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2021, pp. 1–12.

[136] A. C. Tally, J. Abbott, A. Bochner, S. Das, and C. Nippert-Eng, "What mid-career professionals think, know, and feel about phishing: Opportunities for university it departments to better empower employees in their anti-phishing decisions," *Proceedings of the ACM on Human-Computer Interaction*, vol. 7, no. CSCW1, pp. 1–27, 2023.

[137] P. J. Boczkowski, "Mutual shaping of users and technologies in a national virtual community," *Journal of Communication*, vol. 49, no. 2, pp. 86–108, 1999.

[138] O. Sarker, S. Haggag, A. Jayatilaka, and C. Liu, "Personalized guidelines for design, implementation and evaluation of anti-phishing interventions," in *International Symposium on Empirical Software Engineering and Measurement (ESEM)*, 2023.

[139] O. Sarker, S. Haggag, A. Jayatilaka, and C. Liu, "Personalized guidelines for design, implementation and evaluation of anti-phishing interventions," in *2023 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, IEEE Computer Society, 2023, pp. 1–12.

[140] C. M. Gray, E. Stolterman, and M. A. Siegel, "Reprioritizing the relationship between hci research and practice: Bubble-up and trickle-down effects," in *Proceedings of the 2014 conference on Designing interactive systems*, 2014, pp. 725–734.

[141] J. C. Schweitzer, "How academics and practitioners rate academic research.," 1985.

[142] C. Remy, S. Gegenbauer, and E. M. Huang, "Bridging the theory-practice gap: Lessons and challenges of applying the attachment framework for sustainable hci design," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 1305–1314.

[143] H. Hörig, E. Marincola, and F. M. Marincola, "Obstacles and opportunities in translational research," *Nature medicine*, vol. 11, no. 7, pp. 705–708, 2005.

[144] C. Argyris, "Double loop learning in organizations," *Harvard business review*, vol. 55, no. 5, pp. 115–125, 1977.

[145] L. Reinfelder, R. Landwirth, and Z. Benenson, "Security managers are not the enemy either," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–7.

[146] E. Buie, S. Dray, K. Instone, J. Jain, G. Lindgaard, and A. Lund, "How to bring hci research and practice closer together," in *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, 2010, pp. 3181–3184.

[147] *Phishguide early prototype*. [Online]. Available: https://forms.gle/mezz5YYuBPPCFXEW9.

[148] A. J. Ko, T. D. LaToza, and M. M. Burnett, "A practical guide to controlled experiments of software engineering tools with human participants," *Empirical Software Engineering*, vol. 20, pp. 110–141, 2015.

[149] *Phishguide updated version*. [Online]. Available: https://forms.gle/61ZQsm22Dfspu9Ri7.

[150] H. Khalajzadeh, M. Shahin, H. O. Obie, P. Agrawal, and J. Grundy, "Supporting developers in addressing human-centric issues in mobile apps," *IEEE Transactions on Software Engineering*, vol. 49, no. 4, pp. 2149–2168, 2022.

[151] R. Likert, "A technique for the measurement of attitudes.," *Archives of psychology*, 1932.

[152] K. A. Dawood, K. Y. Sharif, A. A. Ghani, H. Zulzalil, A. Zaidan, and B. Zaidan, "Towards a unified criteria model for usability evaluation in the context of open source software based on a fuzzy delphi method," *Information and Software Technology*, vol. 130, p. 106 453, 2021.

[153] J Kirakowski, M Corbett, and M Sumi, "The software usability measurement inventory," *Br J Educ Technol*, vol. 24, no. 3, pp. 210–2, 1993.

[154] I. DIS, "9241-210: 2010. ergonomics of human system interaction-part 210: Human-centred design for interactive systems," *International Standardization Organization (ISO). Switzerland*, 2009.

[155] O. A. Rotaru, S. Vert, R. Vasiu, and D. Andone, "Standardised questionnaires in usability evaluation. applying standardised usability questionnaires in digital products evaluation," in *International Conference on Information and Software Technologies*, Springer, 2020, pp. 39–48.

[156] J. G. Geer, "Do open-ended questions measure "salient" issues?" *Public Opinion Quarterly*, vol. 55, no. 3, pp. 360–370, 1991.

[157] X. Ferré, N. Juristo, H. Windl, and L. Constantine, "Usability basics for software developers," *IEEE software*, vol. 18, no. 1, pp. 22–29, 2001.

[158] A. Strauss and J. Corbin, "Basics of qualitative research techniques," 1998.

[159] D. Bertram, "Likert scales," *Retrieved November*, vol. 2, no. 10, pp. 1–10, 2007.

[160] M. Hennink and B. N. Kaiser, "Sample sizes for saturation in qualitative research: A systematic review of empirical tests," *Social science & medicine*, vol. 292, p. 114 523, 2022.

[161] Mimecast. [Online]. Available: https://www.mimecast.com/.

[162] W. Yeoh, H. Huang, W.-S. Lee, F. Al Jafari, and R. Mansson, "Simulated phishing attack and embedded training campaign," *Journal of Computer Information Systems*, vol. 62, no. 4, pp. 802–821, 2022.

[163] J. Hielscher, U. Menges, S. Parkin, A. Kluge, and M. A. Sasse, ""employees who don't accept the time security takes are not aware enough": The ciso view of human-centred security," in *32st USENIX Security Symposium (USENIX Security 23), Boston, MA*, 2023.

[164] L. Gamisch and D. Pöhn, "A study of different awareness campaigns in a company," in *Proceedings of the 18th International Conference on Availability, Reliability and Security*, 2023, pp. 1–8.

[165] S. I. Hashmi, N. George, E. Saqib, F. Ali, N. Siddique, S. Kashif, S. Ali, N. U. H. Bajwa, and M. Javed, "Training users to recognize persuasion techniques in vishing calls," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–8.

[166] Phishtank. [Online]. Available: https://phishtank.org/.

[167] B. Flyvbjerg, "Five misunderstandings about case-study research," *Qualitative inquiry*, vol. 12, no. 2, pp. 219–245, 2006.

[168] A. Chidukwani, S. Zander, and P. Koutsakis, "A survey on the cyber security of small-to-medium businesses: Challenges, research focus and recommendations," *IEEE Access*, vol. 10, pp. 85 701–85 719, 2022.

[169] ACSC, *Australian cyber security center*. [Online]. Available: https://www.cyber.gov.au/.

[170] C. Onwubiko and A. P. Lenaghan, "Managing security threats and vulnerabilities for small to medium enterprises," in *2007 IEEE Intelligence and Security Informatics*, IEEE, 2007, pp. 244–249.

[171] Z. T. Sworna, C. Islam, and M. A. Babar, "Apiro: A framework for automated security tools api recommendation," *ACM Transactions on Software Engineering and Methodology*, vol. 32, no. 1, pp. 1–42, 2023.