e-JURNAL PENYELIDIKAN DAN INOVASI
ISLAMIC INNOVATION CENTER FOR TECHNOLOGY, UIS
https://ejpi.uis.edu.my/index.php/ejpi
e-ISSN 2289-7909

VOL. 13 NO. 2 (AUGUST, 2026) Special Edition AIDI
Page 218-231
DOI: https://doi.org/10.53840/e-jpi.v13i2.426

Received: 5 August 2026 Accepted: 18 August 2026 Published: 31 August 2026

Understanding Students' Protective Motivation Against AI-Generated Phishing Emails: A Protection Motivation Theory Framework

Memahami Motivasi Perlindungan Pelajar terhadap E-mel Pancingan Data Janaan AI: Kerangka Teori Motivasi Perlindungan

Hiba Khalid Tahir[1], Mohd Fared Abdul Khir[2]
[1] Faculty of Science and Technology, Universiti Sains Islam Malaysia
[2] CyberSecurity and Systems Research Unit (CSS) Faculty of Science and Technology, Universiti Sains Islam Malaysia
Email: [1]3242765@raudah.usim.edu.my [2]fared.khir@usim.edu.my

ABSTRACT

The increasing use of artificial intelligence has enabled attackers to generate realistic, personalised and contextually relevant phishing emails that are more difficult to distinguish from legitimate communications. Undergraduate students may be particularly exposed to these threats because they rely heavily on email and digital platforms for academic and administrative activities. Therefore, this study proposes a conceptual framework based on Protection Motivation Theory (PMT), which includes perceived severity, perceived vulnerability, response efficacy, self-efficacy and response cost as factors influencing undergraduate students' protection motivation and, consequently, their protective behaviour against AI-generated phishing emails. Protection motivation is further proposed to mediate the relationships between the PMT factors and protective behaviour. The framework provides a theoretical explanation of how threat and coping appraisals may shape students' motivation to adopt protective actions against increasingly sophisticated phishing emails. It also provides a foundation for future empirical research and practical guidance for universities in developing cybersecurity awareness and training initiatives that strengthen students' motivation and ability to adopt protective behaviours.

Keywords: AI-generated phishing emails, Protection Motivation Theory, protection motivation, protective behaviour, undergraduate students

INTRODUCTION

The development of generative artificial intelligence (AI) has created new possibilities for producing realistic and personalised content that may be used in social engineering and phishing attacks (Schmitt & Flechais, 2024). AI-generated phishing emails can contain grammatically accurate language and

varied content, which may make some of them more convincing than traditional phishing messages containing obvious errors (Eze & Shamir, 2024).
Large language models can also generate personalised phishing emails using a small amount of information about a recipient, reducing some of the time and cost associated with developing targeted messages (Heiding et al., 2024). These capabilities may increase the realism, personalisation and scale of phishing attempts (Jabir et al., 2025).

Phishing is relevant to higher education because universities hold personal and institutional information and include large populations that regularly use digital services (Casagrande et al., 2023). Undergraduate students commonly use email and online platforms to access academic information and communicate with their institutions, creating opportunities for phishing messages to imitate familiar university communication (Nyasvisvo & Chigada, 2023). Such messages may attempt to persuade recipients to disclose sensitive information, open malicious attachments or visit fraudulent websites (Nyasvisvo & Chigada, 2023). Studies conducted in university settings have recorded interactions with simulated phishing emails among members of university communities, including students (Casagrande et al., 2023; Díaz Vivas et al., 2024). These findings do not mean that all students are highly susceptible, but they indicate that phishing remains relevant when examining cybersecurity behaviour in higher education.

Technical security measures can assist in detecting and blocking phishing emails before they reach users (Zhuo et al., 2023). Nevertheless, users may still encounter phishing messages that are not identified by these controls (Zhuo et al., 2023). In such situations, the outcome may be influenced by how users evaluate the message and whether they perform an appropriate protective action (Desolda et al., 2022). These actions may include verifying the sender, examining links, avoiding suspicious attachments and reporting questionable messages. Examining the factors associated with these actions may complement existing technical approaches to phishing prevention.

Previous research has examined phishing-detection technologies, message characteristics and factors associated with users' susceptibility to phishing (Desolda et al., 2022; Zhuo et al., 2023). More recent studies have considered how generative AI may affect the development and characteristics of phishing emails (Eze & Shamir, 2024; Heiding et al., 2024; Schmitt & Flechais, 2024). However, research specifically examining human factors in generative-AI phishing remains an emerging area (Jabir et al., 2025). Further theoretical attention may therefore help explain how undergraduate students evaluate AI-generated phishing threats and what may motivate them to adopt protective behaviour. This focus moves beyond the recognition of suspicious messages by also considering the motivation to perform protective actions.

Protection Motivation Theory provides a relevant perspective for examining how individuals evaluate threats and possible protective responses Its focus on threat appraisal, coping appraisal and protection motivation makes it applicable to the examination of students' responses to AI-generated phishing emails (Haag et al., 2021).

Accordingly, this study proposes a conceptual framework based on PMT to examine undergraduate students' protective behaviour against AI-generated phishing emails. The framework positions perceived severity, perceived vulnerability, response efficacy, self-efficacy and response cost as factors that may influence protection motivation, which may subsequently influence protective behaviour. Protection motivation is also proposed as a potential mediator between the PMT factors and protective behaviour. By applying PMT to this context, the study offers a theoretical basis for examining students' responses to AI-generated phishing emails. The proposed framework may also support future empirical research and inform the development of cybersecurity awareness and training initiatives in higher education.

LITERATURE REVIEW

AI-Generated Phishing Emails

Generative AI has introduced new capabilities for producing sophisticated phishing emails (Eze & Shamir, 2024). Traditional phishing emails have often relied on generic templates and may contain grammatical or spelling errors that alert recipients to possible deception (Alabdan, 2020).

In comparison, AI-generated phishing emails can contain coherent, grammatically accurate and contextually appropriate language that resembles legitimate communication (Tooher & Lallie, 2025). However, this does not mean that every AI-generated phishing email is more convincing than a traditional one, as its effectiveness may depend on the quality of the generated content and the recipient's response. Advances in generative AI can also support greater personalisation.

Generative AI may imitate writing styles, organisational communication patterns and professional language to produce human-like messages (Schmitt & Flechais, 2024). Publicly available information may also be incorporated into prompts to generate messages tailored to particular recipients and situations (Arif et al., 2024). These capabilities allow phishing emails to reflect familiar communication contexts rather than relying entirely on general messages.

Another characteristic of AI-generated phishing is the automation of content generation. Generative AI can reduce the manual effort required to prepare phishing messages and enable the rapid production of customised content (Hazell, 2023; Kumar et al., 2024). Heiding et al. (2024) demonstrated that large language models could generate personalised phishing emails and reported differences in user responses to AI-generated, manually created and generic phishing messages. Jabir et al. (2025) identified similar scalability and reduced technical requirements as relevant characteristics of generative-AI phishing.

Therefore, generative AI may enhance phishing through linguistic accuracy, personalisation and automated content generation. These capabilities can increase the variety and scale of phishing emails, although their success continues to depend on how recipients assess and respond to the messages. The implications of these developments are particularly relevant within higher education, where students and staff regularly receive communication from numerous academic and administrative sources.

Phishing Attacks in Higher Education

Higher education institutions increasingly rely on institutional email, learning management systems and online platforms to support teaching, research and administrative activities (Lallie et al., 2023). Although these technologies improve access to university services, their extensive use also exposes institutions and users to cybersecurity risks (Ulven & Wangen, 2021). Among these risks, phishing remains a relevant concern within higher education environments (Casagrande et al., 2023).

Phishing is a social engineering attack in which cybercriminals impersonate legitimate organisations or trusted individuals to persuade users to disclose sensitive information (Naqvi et al., 2023). A typical phishing attack directs the recipient to a fraudulent website, where submitted login credentials may be used to gain unauthorised access to a legitimate account (Gouda, 2023). The continued development of phishing techniques has consequently attracted attention within cybersecurity research (Kyaw et al., 2024).

Higher education institutions may be targeted because they manage students' personal records, financial information and research-related data (Okeke & Amaechi, 2024). The extensive use of institutional email and the presence of large university populations also create opportunities for phishing messages

to reach students and staff (Morrow, 2024). Such messages may imitate familiar academic or administrative communications, making their legitimacy more difficult to assess (Lallie et al., 2023).

Empirical studies have demonstrated that members of university communities may interact with simulated phishing messages. Díaz Vivas et al. (2024) found that some participants opened simulated phishing emails and interacted with embedded links. Morrow (2024) similarly reported user interaction with phishing emails in a university setting. These findings do not establish that all university students are vulnerable, but they show that phishing can lead to unsafe responses within higher education environments. This context provides a basis for examining the behavioural factors that may influence undergraduate students' responses to emerging threats, particularly AI-generated phishing emails.

Human Factors in Phishing Attacks

Human factors are important in phishing research because users' decisions can influence whether a phishing attempt succeeds (Kavvadias & Kotsilieris, 2025). Although technical safeguards can reduce exposure to malicious messages, attackers may exploit cognitive and psychological characteristics to influence users' responses (Moustafa et al., 2021; Pollini et al., 2022). Phishing messages commonly use trust, authority, urgency, curiosity and fear to encourage recipients to disclose information or perform unsafe actions (Schmitt & Flechais, 2024).

Generative AI may strengthen these techniques by enabling the production of personalised, contextually relevant and grammatically accurate messages (Jabir et al., 2025). Malloy et al. (2025) found that participants experienced greater difficulty identifying phishing messages created through a combination of human input and generative AI than those produced solely by humans or AI.

When phishing emails reach users' inboxes, their responses may be shaped by knowledge, threat perception, self-efficacy, security awareness and situational influences (Asfoor et al., 2020; Zhuo et al., 2023). Cybersecurity education and training may strengthen users' ability to evaluate suspicious messages and respond safely (Jampen et al., 2020). Therefore, understanding these psychological and behavioural factors is important when examining protective responses to AI-generated phishing emails. This provides a basis for applying Protection Motivation Theory in the proposed conceptual framework.

Protection Motivation Theory

Protection Motivation Theory originally proposed by Rogers (1975) and later revised by Rogers (1983), explains how individuals decide whether to adopt protective behaviours when confronted with potential threats (Ng et al., 2021). The theory was initially developed in health psychology to explain how fear appeals influence behavioural change (Marikyan & Papagiannidis, 2023). It has subsequently been applied in other risk-related contexts, including information systems and cybersecurity (Haag et al., 2021).PMT proposes that individuals cognitively evaluate information about a threat and the available protective responses before deciding whether to act (Mou et al., 2022). These evaluations help explain why individuals exposed to similar threats may develop different levels of protection motivation and demonstrate different protective behaviours (Lowry et al., 2023).

In cybersecurity research, PMT has been used to examine intentions and protective behaviours associated with threats such as phishing, password security and information security compliance (Haag et al., 2021; Kiran et al., 2025). Responses to phishing emails may depend on how individuals perceive the threat and evaluate their ability to perform protective actions (Diman & Abdul Rahman, 2023). PMT therefore provides a relevant theoretical foundation for examining undergraduate students' protective behaviour against AI-generated phishing emails. The following sections explain the two appraisal processes of PMT and their application in the proposed framework.

PMT Components

PMT explains protective behaviour through two cognitive processes: threat appraisal and coping appraisal (Rogers, 1983). Through these processes, individuals evaluate a potential threat and the

available protective responses before developing protection motivation. Protection motivation is therefore shaped by the combined outcomes of threat appraisal and coping appraisal (Lahiri et al., 2021).

Threat Appraisal

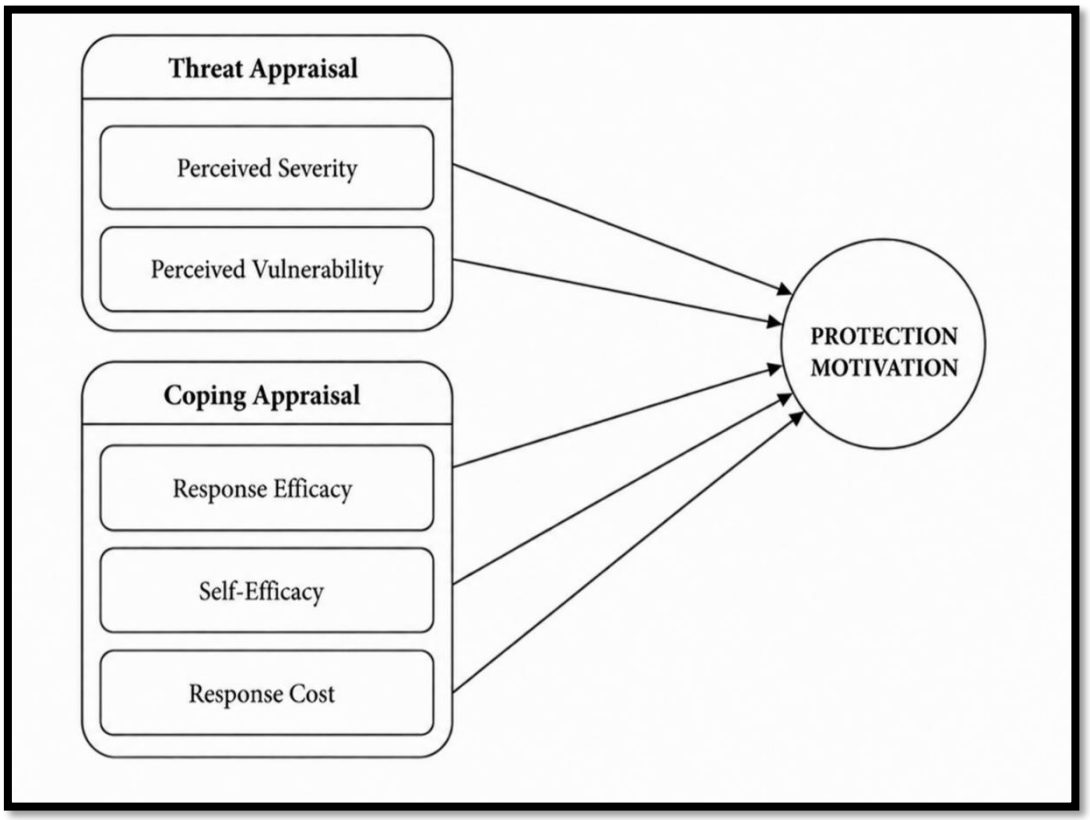
Threat appraisal refers to an individual's evaluation of a potential threat and comprises perceived severity and perceived vulnerability (Lahiri et al., 2021). Perceived severity concerns the seriousness of the possible consequences, while perceived vulnerability reflects the perceived likelihood of being affected by the threat. Together, these evaluations influence whether individuals recognise a need for protective action (Hedayati et al., 2023).

Individuals may be more inclined to consider protective action when they perceive a threat as serious and personally relevant (Lahiri et al., 2021). However, recognising a threat does not necessarily lead to protective behaviour. Individuals may remain unwilling to act if they believe that the available responses are ineffective or difficult to perform (Zou et al., 2023). Threat appraisal should therefore be considered alongside coping appraisal when explaining protective behaviour.

Coping Appraisal

Coping appraisal refers to an individual's evaluation of the available protective responses and comprises response efficacy, self-efficacy and response cost (Mou et al., 2022). It considers whether a recommended protective measure is believed to be effective, whether the individual feels capable of performing it and whether its associated costs are acceptable (Dodge et al., 2023).

While threat appraisal concerns the importance and personal relevance of a threat, coping appraisal concerns whether protective action is viewed as effective and achievable (Mou et al., 2022). Individuals may be more willing to adopt protective behaviour when they believe that the recommended response can reduce the threat, feel confident in performing it and perceive limited barriers to taking action (Dodge et al., 2023). In cybersecurity contexts, coping appraisal has been associated with intentions to adopt protective responses to cyber threats, including phishing (Naqvi et al., 2023). Together, threat appraisal and coping appraisal explain how individuals evaluate threats and develop motivation to perform protective behaviour (Li et al., 2022). Figure 1 illustrates the two appraisal processes of PMT and their role in shaping protection motivation and protective behaviour.



Source: Based on Rogers (1983)

Figure 1: Protection Motivation Theory

Proposed Conceptual Framework

This paper proposes a conceptual framework for explaining undergraduate students' protective behaviour against AI-generated phishing emails. Based on PMT, the framework suggests that perceived severity, perceived vulnerability, response efficacy, self-efficacy and response cost shape protection motivation, which subsequently influences protective behaviour. Protection motivation is also positioned as the mechanism connecting the five PMT factors with protective behaviour. Figure 2 presents the proposed conceptual framework.

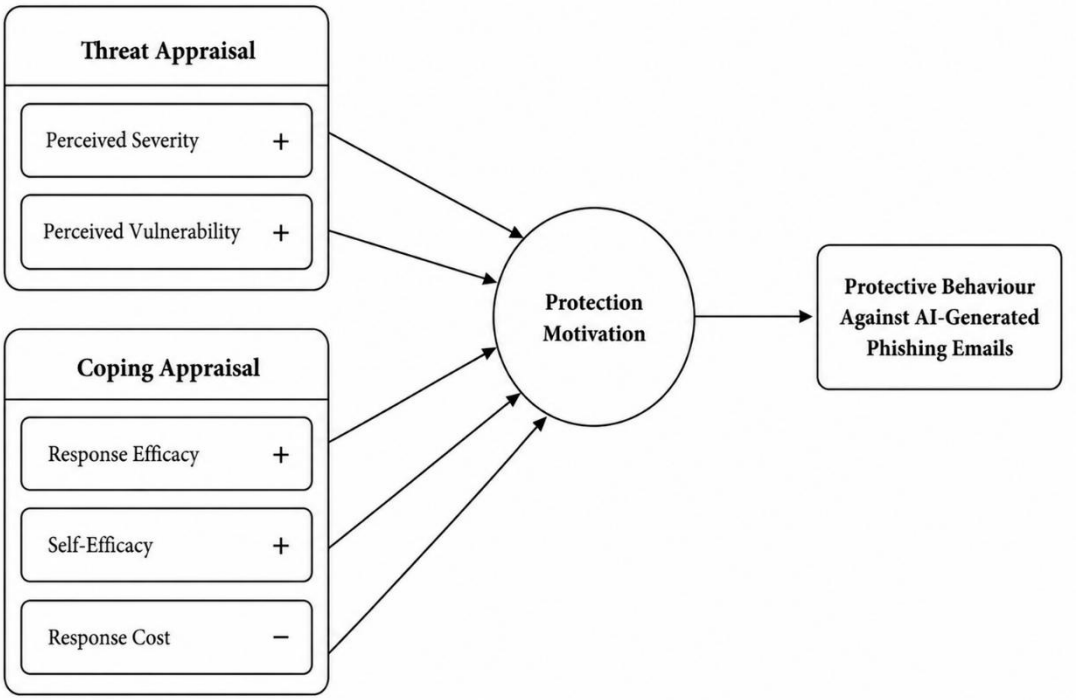


Figure 2: Proposed Conceptual Framework

PMT Factors and Protection Motivation

Threat appraisal in the proposed framework comprises perceived severity and perceived vulnerability. Perceived severity refers to an individual's assessment of the seriousness of a threat and its potential consequences (Kuo et al., 2025). In the context of AI-generated phishing emails, students may consider possible consequences such as financial loss, identity theft, privacy breaches and unauthorised access to personal or academic information (Gwenhure, 2025). When these consequences are perceived as serious, students may develop greater motivation to adopt appropriate protective measures (Han et al., 2025; Kiran et al., 2025).

Perceived vulnerability concerns an individual's assessment of the likelihood of personally experiencing a threat (Balla & Hagger, 2025). AI-generated phishing emails may increase this concern because their realistic and personalised content can make them difficult to distinguish from legitimate communication (Okokpujie et al., 2025). Students who regard themselves as potential targets may therefore be more motivated to adopt protective measures against these emails (Gwenhure, 2025; Pratama et al., 2025).

Coping appraisal comprises response efficacy, self-efficacy and response cost. Response efficacy refers to the belief that recommended protective measures can prevent or reduce the effects of a threat. For AI-generated phishing emails, these measures may include verifying the sender's identity, avoiding suspicious links or attachments and reporting suspicious messages (Alam et al., 2025; Suzuki & Salinas Monroy, 2021). When students believe that these measures are effective, they may be more motivated to apply them when evaluating potentially malicious emails (Han et al., 2025; Vortia, 2025).

.

Self-efficacy refers to an individual's confidence in their ability to perform the recommended protective behaviour (Lee et al., 2023). In this context, it reflects students' confidence in recognising, evaluating and responding appropriately to AI-generated phishing emails. Students who believe that they possess

the necessary knowledge and skills may develop stronger protection motivation and be more willing to apply appropriate cybersecurity practices (Kim, 2025; Vafaei-Zadeh et al., 2024).

Response cost represents the perceived time, effort, inconvenience or resources required to perform a protective action (Balla & Hagger, 2025). Practices such as verifying suspicious emails, enabling multi-factor authentication or following additional authentication procedures may be perceived as inconvenient or time-consuming (Bax et al., 2021). When these costs are considered excessive, students' motivation to adopt the recommended protective measures may be reduced (Han et al., 2025; Kiran et al., 2025). Response cost is therefore included in the proposed framework to account for the potential barriers students may consider when deciding whether to adopt protective measures against AI-generated phishing emails.

Protection Motivation and Protective Behaviour

Protection motivation refers to an individual's intention or willingness to adopt actions that reduce exposure to a potential threat (Carter & McNealey, 2025). In the proposed framework, it represents students' willingness to protect themselves after evaluating the possible consequences of AI-generated phishing emails and the available protective responses.

Protective behaviour refers to the actions performed to avoid or reduce phishing-related risks. These actions may include verifying the authenticity of emails, avoiding suspicious links and attachments, using multi-factor authentication, updating security settings and reporting suspicious messages (Jaeger & Eckhardt, 2021). Previous cybersecurity research indicates that individuals with stronger protection motivation are generally more likely to apply recommended security practices and avoid risky online behaviour (Kiran et al., 2025; Van 't Hoff-de Goede et al., 2025). Accordingly, the proposed framework suggests that undergraduate students' protection motivation may encourage protective behaviour against AI-generated phishing emails.

Mediating Role of Protection Motivation

Awareness of a cybersecurity threat may not necessarily lead directly to protective behaviour because individuals must also be willing to apply the recommended security measures (Zwilling et al., 2022). Within PMT, protection motivation develops from individuals' evaluations of the threat and their ability to cope with it. It may therefore explain how the five PMT factors are translated into protective behaviour.

Previous studies have identified protection motivation as an important predictor of information-security behaviour (Alrawhani et al., 2025; Mou et al., 2022). In phishing contexts, it may connect individuals' assessments of threat severity, personal vulnerability, response effectiveness, personal capability and response costs with their subsequent protective actions (Bax et al., 2021). Accordingly, the proposed framework positions protection motivation as a mediator between the five PMT factors and undergraduate students' protective behaviour against AI-generated phishing emails.

Theoretical and Practical Implications

The proposed conceptual framework offers theoretical and practical implications for understanding undergraduate students' protective behaviour against AI-generated phishing emails. As the framework has not yet been empirically tested, these implications represent potential contributions that should be examined in future research.

Theoretical Implications

The proposed framework extends the application of Protection Motivation Theory to the emerging context of AI-generated phishing emails among undergraduate students. Although PMT has been widely applied to cybersecurity behaviour (Haag et al., 2021; Kiran et al., 2025).

The increasing realism, personalisation and linguistic accuracy of AI-generated phishing emails create a developing threat context that requires further behavioural investigation (Heiding et al., 2024; Schmitt & Flechais, 2024).

By integrating perceived severity, perceived vulnerability, response efficacy, self-efficacy and response cost, the framework provides a structured explanation of how students' threat and coping appraisals may shape their protection motivation. It further positions protection motivation as the mechanism through which these cognitive assessments may influence protective behaviour. This proposed mediating role offers a clearer explanation of how students' perceptions may be translated into actions such as verifying senders, avoiding suspicious links and reporting suspicious emails. The framework therefore provides a theoretical basis for future empirical research on protective responses to AI-generated phishing emails.

Practical Implications

The proposed framework may assist universities in designing cybersecurity initiatives that address both students' perceptions of phishing threats and their ability to respond effectively. Awareness programmes could explain the possible consequences of AI-generated phishing attacks and students' potential exposure to them. However, such programmes should avoid exaggerated fear-based messages and provide realistic information about the threat.

Universities could also strengthen response efficacy and self-efficacy by demonstrating practical measures for identifying and managing suspicious emails. Training activities may include recognising unusual requests, verifying senders through trusted channels, avoiding unverified links and attachments, using multi-factor authentication and reporting suspicious messages. Phishing simulations may also provide students with opportunities to practise these measures in realistic situations (Jampen et al., 2020).

Universities should additionally consider response cost when developing cybersecurity procedures. Protective measures that are perceived as complicated, inconvenient or time-consuming may discourage consistent adoption (Bax et al., 2021). Therefore, reporting systems and verification procedures should be clear and accessible. These implications remain conceptual and require empirical evaluation before firm recommendations can be made.

Future Research

Future research should empirically examine the proposed conceptual framework to determine how perceived severity, perceived vulnerability, response efficacy, self-efficacy and response cost influence undergraduate students' protection motivation and protective behaviour against AI-generated phishing emails. Particular attention should be given to the proposed mediating role of protection motivation in explaining how students' threat and coping appraisals are translated into protective actions.

Studies involving undergraduate students from different universities and educational contexts could provide a broader understanding of whether these relationships vary across institutions or student populations. Researchers may also compare students' responses to AI-generated and traditional phishing emails to determine whether the characteristics of AI-generated messages influence their threat perceptions and protective responses differently.

Future studies could combine questionnaires with phishing simulations or controlled experiments. This would allow researchers to compare students' reported intentions with their actual responses to suspicious emails. Longitudinal research may also help explain whether students' protection motivation and protective behaviour change over time as AI-generated phishing techniques and cybersecurity awareness programmes continue to develop. These investigations would provide empirical evidence for assessing and refining the proposed framework.

CONCLUSION

Generative AI has increased the sophistication of phishing emails by enabling attackers to produce realistic, personalised and grammatically accurate messages. This development may make phishing attempts more difficult for undergraduate students to recognise, highlighting the need to understand the

factors that encourage protective responses.This paper proposed a PMT framework in which perceived severity, perceived vulnerability, response efficacy, self-efficacy and response cost influence protection motivation, which subsequently shapes protective behaviour. Protection motivation is also proposed as the mechanism through which these factors translate into protective action.The framework extends the application of PMT to AI-generated phishing in higher education and provides a foundation for developing more targeted cybersecurity awareness and training initiatives. Future empirical research should test and refine the proposed relationships across different higher education settings.

REFERENCES

Alabdan, R. (2020). Phishing attacks survey: Types, vectors, and technical approaches. Future Internet, 12(10), Article 168. https://doi.org/10.3390/fi12100168

Alam, S. S., Ahsan, N., Kokash, H. A., Alam, S., & Ahmed, S. (2025). A students' behaviors in information security: Extension of protection motivation theory (PMT). Information Security Journal: A Global Perspective, 34(2), 191–213. https://doi.org/10.1080/19393555.2024.2408264

Alrawhani, E. M., Romli, A., & Al-Sharafi, M. A. (2025). Evaluating the role of protection motivation theory in information security policy compliance: Insights from the banking sector using PLS-SEM approach. Journal of Open Innovation: Technology, Market, and Complexity, 11(1), Article 100463. https://doi.org/10.1016/j.joitmc.2024.100463

Arif, A., Khan, M. I., & Khan, A. R. A. (2024). An overview of cyber threats generated by AI. International Journal of Multidisciplinary Sciences and Arts, 3(4), 67–76. https://doi.org/10.47709/ijmdsa.v3i4.4753

Asfoor, A. H., Rahim, F. A., & Yussof, S. (2020). Identifying factors that influence security behaviors relating to phishing attacks susceptibility: A systematic literature review. Journal of Theoretical and Applied Information Technology, 98(15), 3127–3161.

Balla, J., & Hagger, M. S. (2025). Protection motivation theory and health behaviour: Conceptual review, discussion of limitations, and recommendations for best practice and future research. Health Psychology Review, 19(1), 145–171. https://doi.org/10.1080/17437199.2024.2413011

Bax, S., McGill, T., & Hobbs, V. (2021). Maladaptive behaviour in response to email phishing threats: The roles of rewards and response costs. Computers & Security, 106, Article 102278. https://doi.org/10.1016/j.cose.2021.102278

Carter, T., & McNealey, R. L. (2025). Protection motivation and cybersecurity intentions: A visual conjoint experiment. Journal of Experimental Criminology. Advance online publication. https://doi.org/10.1007/s11292-025-09717-1

Casagrande, M., Conti, M., Fedeli, M., & Losiouk, E. (2023). Alpha Phi-shing fraternity: Phishing assessment in a higher education institution. Journal of Cybersecurity Education, Research and Practice, 2022(2), Article 2. https://doi.org/10.32727/8.2023.1

Desolda, G., Ferro, L. S., Marrella, A., Catarci, T., & Costabile, M. F. (2022). Human factors in phishing attacks: A systematic literature review. ACM Computing Surveys, 54(8), Article 173, 1–35. https://doi.org/10.1145/3469886

Díaz Vivas, D. E., Gutierrez Pena, W. Y., Cristancho Botero, S. P., & Rojas, A. E. (2024). A controlled phishing attack in a university community: A case study. Journal of Internet Services and Information Security, 14(2), 98–110. https://doi.org/10.58346/JISIS.2024.I2.007

Diman, A. P., & Abdul Rahman, T. K. (2023). Examining individual tendency to respond to phishing e-mails from the perspective of protection motivation theory. Journal of Education and Social Sciences, 25(1), 37–50.

Dodge, C. E., Fisk, N., Burruss, G. W., & Jaynes, C. M. (2023). What motivates users to adopt cybersecurity practices? A survey experiment assessing protection motivation theory. Criminology & Public Policy. Advance online publication. https://doi.org/10.1111/1745-9133.12641

Eze, C. S., & Shamir, L. (2024). Analysis and prevention of AI-based phishing email attacks. Electronics, 13(10), Article 1839. https://doi.org/10.3390/electronics13101839

Gouda, S. K. (2023, April). A comprehensive study of phishing attacks and their countermeasures [Preprint]. ResearchGate. https://doi.org/10.13140/RG.2.2.36686.13120

Gwenhure, A. K. (2025). University students' security behavior against email phishing attacks: Insights from the health belief model. Journal of Cybersecurity, 11(1), Article tyaf034. https://doi.org/10.1093/cybsec/tyaf034

Haag, S., Siponen, M., & Liu, F. (2021). Protection motivation theory in information systems security research: A review of the past and a road map for the future. The DATA BASE for Advances in Information Systems, 52(2), 25–67. https://doi.org/10.1145/3462766.3462770

Han, M., Zhao, H., Ma, X., & Shi, R. (2025). Influencing factors of information security behavior among college students based on protection motivation theory: Evidence from China. Frontiers in Public Health, 13, Article 1677024. https://doi.org/10.3389/fpubh.2025.1677024

Hazell, J. (2023). Spear phishing with large language models [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2305.06972

Hedayati, S., Damghanian, H., Farhadinejad, M., & Rastgar, A. A. (2023). Meta-analysis on application of protection motivation theory in preventive behaviors against COVID-19. International Journal of Disaster Risk Reduction, 94, Article 103758. https://doi.org/10.1016/j.ijdrr.2023.103758

Heiding, F., Schneier, B., Vishwanath, A., Bernstein, J., & Park, P. S. (2024). Devising and detecting phishing emails using large language models. IEEE Access, 12, 42131–42146. https://doi.org/10.1109/ACCESS.2024.3375882

Jabir, R., Le, J., & Nguyen, C. (2025). Phishing attacks in the age of generative artificial intelligence: A systematic review of human factors. AI, 6(8), Article 174. https://doi.org/10.3390/ai6080174

Jaeger, L., & Eckhardt, A. (2021). Eyes wide open: The role of situational information security awareness for security-related behaviour. Information Systems Journal, 31(3), 429–472. https://doi.org/10.1111/isj.12317

Jampen, D., Gür, G., Sutter, T., & Tellenbach, B. (2020). Don't click: Towards an effective anti-phishing training. A comparative literature review. Human-Centric Computing and Information Sciences, 10, Article 33. https://doi.org/10.1186/s13673-020-00237-7

Kavvadias, A., & Kotsilieris, T. (2025). Understanding the role of demographic and psychological factors in users' susceptibility to phishing emails: A review. Applied Sciences, 15(4), Article 2236. https://doi.org/10.3390/app15042236

Kim, S. (2025). Impact of cybersecurity self-efficacy on digital economic behaviors among older adults. INQUIRY: The Journal of Health Care Organization, Provision, and Financing, 62, Article 469580251370933. https://doi.org/10.1177/00469580251370933

Kiran, U., Khan, N. F., Murtaza, H., Farooq, A., & Pirkkalainen, H. (2025). Explanatory and predictive modeling of cybersecurity behaviors using protection motivation theory. Computers & Security, 149, Article 104204. https://doi.org/10.1016/j.cose.2024.104204

Kumar, S., Menezes, A., Giri, S., & Kotikela, S. (2024). What the phish! Effects of AI on phishing attacks and defense. International Conference on AI Research, 4(1), 218–226. https://doi.org/10.34190/icair.4.1.3224

Kuo, S.-H., Lin, H.-M., & Chiang, H.-C. (2025). Perceived severity, anxiety, and protection motivation in shaping protection insurance product purchase intentions: Evidence from the COVID-19 public health crises. Journal of Risk and Financial Management, 18(12), Article 722. https://doi.org/10.3390/jrfm18120722

Kyaw, P. H., Gutierrez, J., & Ghobakhlou, A. (2024). A systematic review of deep learning techniques for phishing email detection. Electronics, 13(19), Article 3823. https://doi.org/10.3390/electronics13193823

Lahiri, A., Jha, S. S., Chakraborty, A., Dobe, M., & Dey, A. (2021). Role of threat and coping appraisal in protection motivation for adoption of preventive behavior during COVID-19 pandemic. Frontiers in Public Health, 9, Article 678566. https://doi.org/10.3389/fpubh.2021.678566

Lallie, H. S., Thompson, A., Titis, E., & Stephens, P. (2023). Understanding cyber threats against universities, colleges, and schools [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2307.07755

Lee, Y. Y., Gan, C. L., & Liew, T. W. (2023). Thwarting instant messaging phishing attacks: The role of self-efficacy and the mediating effect of attitude towards online sharing of personal information. International Journal of Environmental Research and Public Health, 20(4), Article 3514. https://doi.org/10.3390/ijerph20043514

Li, L., Xu, L., & He, W. (2022). The effects of antecedents and mediating factors on cybersecurity protection behavior. Computers in Human Behavior Reports, 5, Article 100165. https://doi.org/10.1016/j.chbr.2021.100165

Lowry, P. B., Moody, G. D., Parameswaran, S., & Brown, N. J. (2023). Examining the differential effectiveness of fear appeals in information security management using two-stage meta-analysis. Journal of Management Information Systems, 40(4), 1099–1138. https://doi.org/10.1080/07421222.2023.2267318

Malloy, T., Ferreira, M. J., Fang, F., & Gonzalez, C. (2025). Training users against human and GPT-4 generated social engineering attacks. In A. Moallem (Ed.), HCI for cybersecurity, privacy and trust (Lecture Notes in Computer Science, Vol. 15814, pp. 54–74). Springer. https://doi.org/10.1007/978-3-031-92833-8_4

Marikyan, D., & Papagiannidis, S. (2023). Protection motivation theory: A review. In S. Papagiannidis (Ed.), TheoryHub book (pp. 78–93). Newcastle University. https://open.ncl.ac.uk/theory-library/TheoryHubBook.pdf

Morrow, E. (2024). Scamming higher ed: An analysis of phishing content and trends. Computers in Human Behavior, 158, Article 108274. https://doi.org/10.1016/j.chb.2024.108274

Mou, J., Cohen, J. F., Bhattacherjee, A., & Kim, J. (2022). A test of protection motivation theory in the information security literature: A meta-analytic structural equation modeling approach. Journal of the Association for Information Systems, 23(1), 196–236. https://doi.org/10.17705/1jais.00723

Moustafa, A. A., Bello, A., & Maurushat, A. (2021). The role of user behaviour in improving cyber security management. Frontiers in Psychology, 12, Article 561011. https://doi.org/10.3389/fpsyg.2021.561011

Naqvi, B., Perova, K., Farooq, A., Makhdoom, I., Oyedeji, S., & Porras, J. (2023). Mitigation strategies against the phishing attacks: A systematic literature review. Computers & Security, 132, Article 103387. https://doi.org/10.1016/j.cose.2023.103387

Ng, K. C., Zhang, X., Thong, J. Y. L., & Tam, K. Y. (2021). Protecting against threats to information security: An attitudinal ambivalence perspective. Journal of Management Information Systems, 38(3), 732–764. https://doi.org/10.1080/07421222.2021.1962601

Nyasvisvo, B., & Chigada, J. M. (2023). Phishing attacks: A security challenge for university students studying remotely. The African Journal of Information Systems, 15(2), Article 3. https://digitalcommons.kennesaw.edu/ajis/vol15/iss2/3/

Okeke, O. C., & Amaechi, C. E. (2024). Awareness of phishing attacks in institutions of higher learning: A review of types and technical approaches. International Journal of Research and Innovation in Applied Science, 9(10), 309–333. https://doi.org/10.51584/IJRIAS.2024.910031

Okokpujie, K., Ariyo, M. A., Moninuola, F. S., Akanle, M. B., & Okokpujie, I. P. (2025). Evaluating students' vulnerability and awareness to phishing attacks in educational institutions. International Journal of Safety and Security Engineering, 15(3), 621–630. https://doi.org/10.18280/ijsse.150320

Pollini, A., Callari, T. C., Tedeschi, A., Ruscio, D., Save, L., Chiarugi, F., & Guerri, D. (2022). Leveraging human factors in cybersecurity: An integrated methodological approach. Cognition, Technology & Work, 24(2), 371–390. https://doi.org/10.1007/s10111-021-00683-y

Pratama, O., Aladin, R. A., Lim, B., & Sundjaja, A. M. (2025). Determinants of security behavior intention in state-owned enterprises: Applying protection motivation theory to phishing emails. International Journal of Safety and Security Engineering, 15(3), 443–453. https://doi.org/10.18280/ijsse.150304

Rogers, R. W. (1975). A protection motivation theory of fear appeals and attitude change. The Journal of Psychology, 91(1), 93–114. https://doi.org/10.1080/00223980.1975.9915803

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. T. Cacioppo & R. E. Petty (Eds.), Social psychophysiology: A sourcebook (pp. 153–176). Guilford Press.

Schmitt, M., & Flechais, I. (2024). Digital deception: Generative artificial intelligence in social engineering and phishing. Artificial Intelligence Review, 57, Article 324. https://doi.org/10.1007/s10462-024-10973-2

Suzuki, Y. E., & Salinas Monroy, S. A. (2021). Prevention and mitigation measures against phishing emails: A sequential schema model. Security Journal, 35(4), 1162–1182. https://doi.org/10.1057/s41284-021-00318-x

Tooher, P., & Lallie, H. S. (2025). A two-stage deep learning framework for AI-driven phishing email detection based on persuasion principles. Computers, 14(12), Article 523. https://doi.org/10.3390/computers14120523

Ulven, J. B., & Wangen, G. (2021). A systematic review of cybersecurity risks in higher education. Future Internet, 13(2), Article 39. https://doi.org/10.3390/fi13020039

Vafaei-Zadeh, A., Nikbin, D., Teoh, K. Y., & Hanifah, H. (2024). Cybersecurity awareness and fear of cyberattacks among online banking users in Malaysia. International Journal of Bank Marketing, 43(3), 476–505. https://doi.org/10.1108/IJBM-03-2024-0138

van ’t Hoff-de Goede, M. S., Leukfeldt, E. R., van de Weijer, S. G. A., & van der Kleij, R. (2025). Does protection motivation predict self-protective online behaviour? Comparing self-reported and actual online behaviour using a population-based survey experiment. Computers in Human Behavior Reports, 18, Article 100649. https://doi.org/10.1016/j.chbr.2025.100649

Vortia, W. (2025). Modelling cybersecurity awareness, perceived threats and secure online behavioral intentions among Ghanaian university students: A PLS-SEM approach. Magna Scientia Advanced Research and Reviews, 14(2), 96–111. https://doi.org/10.30574/msarr.2025.14.2.0094

Zhuo, S., Biddle, R., Koh, Y. S., Lottridge, D., & Russello, G. (2023). SoK: Human-centered phishing susceptibility. ACM Transactions on Privacy and Security, 26(3), Article 24, 1–27. https://doi.org/10.1145/3575797

Zou, X., Chen, Q., Zhang, Y., & Evans, R. (2023). Predicting COVID-19 vaccination intentions: The roles of threat appraisal, coping appraisal, subjective norms, and negative affect. BMC Public Health, 23, Article 230. https://doi.org/10.1186/s12889-023-15169-x

Zwilling, M., Klien, G., Lesjak, D., Wiechetek, Ł., Cetin, F., & Basim, H. N. (2022). Cyber security awareness, knowledge and behavior: A comparative study. Journal of Computer Information Systems, 62(1), 82–97. https://doi.org/10.1080/08874417.2020.1712269