A Review of Social Engineering Threats in the Age of AI: Understanding Human Vulnerabilities Through Protection Motivation Theory (PMT)

Noor Aishah Zainiar[1*], Siti Nur Edayu Hashim[2]
[1,2] *Department of Information Technology and Communication, Politeknik Sultan Idris Shah*
**Corresponding author's email: naishah@psis.edu.my*

Abstract: Social engineering attacks are one of the most widespread cybersecurity threats today, and it is becoming even more dangerous with the advancements of Artificial Intelligence (AI). In contrast to traditional cybersecurity attacks that target system faults, social engineering uses human weaknesses, making it a serious concern for both individuals and companies. The introduction of AI strategies such as deepfakes, automated phishing, and voice cloning has made these attacks more realistic and harder to detect. Given that cybersecurity risks are on the rise, not much is known about how people see and react to threats that are driven by AI. This study applies Protection Motivation Theory (PMT) to explore how individuals perceive the severity, vulnerability, response efficacy, and self-efficacy of social engineering attacks. Findings show that social engineering attacks that are driven by AI change how people think and make decisions, which makes it more likely that someone will fall victim. The results also show how important it is to educate people about cybersecurity. For example, cybersecurity training programs that focus on PMT parts can help people gain the skills and confidence they need to recognize and prevent these complicated threats. This study contributes to the broader effort of strengthening public and educational outreach to combat the growing threats of rapidly evolving AI-driven social engineering.

Keywords: *Social Engineering, Artificial Intelligence (AI) Threats, Cybersecurity Education, Protection Motivation Theory (PMT), Human Factors in Security*

1.0 INTRODUCTION

In today's digital era, the features of cybersecurity threats have significantly transformed, both in scale and complexity, in response to the widespread access to the internet and rapid advancements in technology. Therefore, protecting important data, digital assets, and people from harmful cybersecurity measures is the highest priority (Farooq et al., 2025; Alghazo et al., 2025; Li et al., 2019; Phan et al., 2025; Vulpe et al., 2024). Li et al. (2019) showed that even though businesses and people use a range of technological barriers, like firewalls, password protection systems, and system tracking tools, these are often not enough to completely protect their data. The biggest weakness remains people, though, because they disregard security rules and do risky things that make defenses less effective.

Although modern security systems are in place, the human factor remains a key vulnerability and is generally viewed as the weakest link in cybersecurity defenses (Almansoori et al., 2023; Khadka & Ullah, 2025; Kiran et al., 2025; Phan et al., 2025; Sulaiman et al., 2022). Almansoori et al. (2023) claim that addressing social and behavioral measurements is essential for enhancing modern cybersecurity strategies. Several studies emphasize the significance of understanding users' behaviors, cognitive patterns, and emotional responses as essential elements of effective cybersecurity strategies. (Sulaiman et al., 2022).

Social engineering is one of the most common, identifiable, and straightforward methods of attack among the diverse range of cybersecurity threats. This term encompasses various techniques employed by cybercriminals to force individuals into revealing sensitive information or executing unintended actions (Asker et al., 2024; Faotu et al., 2024, Manyam, 2022; Phan et al., 2025). Instead of focusing on technical shortcomings, these attacks mostly take advantage of psychological weaknesses. Previous study has shown that a lot of security breaches are linked to or start with social engineering techniques, which often involve building trust and changing how victims think about things (Asker et al., 2024; Manyam, 2022).

2.0 LITERATURE REVIEW

The rapid advancement of Artificial Intelligence (AI) is reshaping numerous sectors, including cybersecurity, by offering transformative capabilities (Akhtar & Rawol, 2024; Al Siam et al., 2025; Vulpe et al., 2024). Despite its benefits, AI also introduces notable cybersecurity risks. It is developing how organizations assess threats, implement security controls, and respond to cybersecurity incidents. On one hand, AI significantly enhances the detection and mitigation of cybersecurity threats (Akhtar & Rawol, 2024; Al Siam et al., 2025; Chouraik, El-Founir, & Taibi, 2024; Dong & Kotenko, 2025; Salem et al., 2024). In contrast, it lets cybercriminals create ever more complex and stealthier cyberattacks. AI possesses a dual function, which can serve as either potential danger in cybersecurity and an essential protective measure. Stronger and more flexible security systems are desperately needed as the increase in attack complexity brought on by AI innovation highlights (Alghazo et al., 2025; Chouraik et al., 2024).

AI is having an immense impact on the types of social engineering attacks that are used and how well they work. Cybercriminals are using AI more and more to make scam links, fake websites, misleading emails, and changed social media posts that target specific people and take advantage of their weaknesses. Using machine learning and prediction analytics, cybercriminals can find possible victims and tailor their messages to them to make manipulation more effective (Manyam, 2022). AI technologies like voice cloning and deep fakes develop social engineering attacks. Additionally, Faotu et al. (2024) note that AI-driven automated social engineering bots are transforming threat actor techniques. This shows how quickly, and effectively advanced AI-driven solutions are needed to find and stop these kinds of threats.

According to Al-Hashem & Saidi (2023), one of the most important things to do to protect yourself against cybersecurity threats is to acknowledge and deal with the human factor. Psychological factors play an important part in what people do and choose to do online, and they possess a significant impact on how they understand and respond to risks. Researchers frequently employ the Protection Motivation Theory (PMT), a prominent framework in cybersecurity and information security research, to enhance their comprehension of these behavioral responses (Achuthan et al., 2025; Alghazo et al., 2025; Al-Harthy et al., 2020; Almansoori et al., 2023; Kiran et al., 2025; Li et al., 2019; Sulaiman et al., 2022; Vafaei-Zadeh et al., 2025).

Faotu et al. (2024) pointed out that it is still very important to understand how people see and respond to these threats, even though social engineering tools have become more complicated and easier to find, especially as AI has grown. More research should be done on the psychological effects of AI-driven social engineering, even though some studies have looked at cybersecurity knowledge and user behavior and used PMT to explain how people make decisions about security. (Achuthan et al., 2025; Alghazo et al., 2025; Al-Harthy et al., 2020; Kiran et al., 2025; Sulaiman et al., 2022; Vafaei-Zadeh et al., 2025). Psychological frameworks like PMT can help align cybersecurity with human cognition and behavior (Al-Hashem & Saidi, 2023). According to Al-Harthy et al. (2020), raising awareness of information security awareness (ISA) is crucial to transforming people from passive users into security-conscious actors and lowering the probability and severity of assaults.

3.0 PROTECTION MOTIVATION THEORY (PMT)

Protection Motivation Theory (PMT) is a well-established theoretical framework frequently working to understand how individuals are concerned in adopting protective behaviors when confronted with perceived threats. Originally developed by Rogers (1983) in the context of health psychology, PMT was aimed to explain how persuasive communication, especially fear appeals, impact behavioral change. The framework was afterwards developed to encompass cognitive processes, allowing it to mature into a more universal decision-making framework applicable to a wide range of risk-related activities beyond the health domain.

3.1 Overview of Protection Motivation Theory

According to PMT, an individual's motivation to protect themselves is influenced by two primary cognitive processes: threat appraisal and coping appraisal. These cognitive evaluations shape a user's protection motivation, which is the intention to either adopt, maintain, or avoid a particular behavior in response to a perceived threat (Alrawhani, Romli, & Al-Sharafi, 2025; Jansen & van Schaik, 2016). Essentially, individuals weigh the perceived risks and potential benefits or costs associated with engaging in protective actions (Fisher, 2024; Kiran et al., 2025). Following these assessments, people may use adaptive coping techniques like proactive threat mitigation or dysfunctional ones like denial, avoidance, or inaction (Song, Lee, & Roh, 2024). Maladaptive responses may increase risk, while adaptive responses are protective.

i. Threat Appraisal

   a. Threat appraisal is the cognitive process through which individuals assess the degree of danger presented by a potential threat or hazardous situation. This appraisal involves evaluating two primary components: the perceived severity of the threat and the perceived vulnerability (Debb & McClellan, 2021; Farooq et al., 2019; Kiran et al., 2025; Schneider & Rahman, 2021; Tsai et al., 2016). Additionally, the full PMT model incorporates the concept of maladaptive rewards, referring to the perceived benefits or incentives that may be gained from engaging in unsafe or risky behaviors despite the known threat, which can weaken motivation to adopt protective measures (Marikyan & Papagiannidis, 2023).

ii. Coping Appraisal

   b. According to Sulaiman et al. (2023), the process of coping entails the evaluation of potential coping strategies or actions that are proposed with the intention of reducing, mitigating, or preventing a potentially dangerous security incident or potential threat. A personal judgment of an individual's capacity to deal with the difficulty of the circumstance is included in this appraisal. According to Floyd, Prentice-Dunn, and Rogers (2000), the process examines the capability to deal with and avoid the hazardous situation that is being threatened.

3.2 Components of Protection Motivation Theory

The core components of PMT are typically grouped into the two appraisal processes:

i. Threat Appraisal Components:

   a. Perceived Severity: The seriousness and potential harm associated with a threat. This is the judgment of the importance of danger and the potential impact of consequences. In the context of cybersecurity, perceived severity can be a judgment of severe damage at work or anywhere with an internet connection.
   b. Perceived Vulnerability: The extent to which a user considers they could be a victim or are prone to be exposed to the threat as well as the personal probability or possibility of a security incident happening. It captures the anxiety of a cyberattack and the awareness one does not have preventative measures.
   c. Maladaptive Rewards: Rewards associated with engaging in risky or unsafe behaviors.

ii. Coping Appraisal Components:

   a. Response Efficacy: The perceived effectiveness of a recommended response or adaptive behavior in reducing or mitigating a threat. In cybersecurity, this is the perception that secure behavior benefits the individual by mitigating threats.

b. Self-Efficacy: Users' certainty in their own competence to effectively execute the advised protective response or adaptive behavior defines their self-efficacy. Believing confidence or faith in one's competence; self-efficacy was included in PMT as a further factor determining protection motive. It is the ability of a person to implement cybersecurity practice.
c. Response Costs: The perceived costs associated with performing the protective behavior. This can include psychological or physical costs and may be financial or temporal.
d. Perceived Cost of Compliance and Perceived Benefits When Complying: These are also distinguished as factors in PMT related to coping appraisal.

The combination of threat appraisal and coping appraisal represents the core foundation of the PMT (Jansen & van Schaik, 2016; Marikyan & Papagiannidis, 2023). At the heart of PMT is protection motivation, which functions as a mediating variable that links cognitive evaluations to behavioral outcomes (Balla & Hagger, 2025). This motivation reflects an individual's intention to adopt protective behaviors in response to a perceived threat. Moreover, Debb & McClellan, 2021 indicated that an individual’s prior knowledge and skill can considerably impact both forms of evaluations, hence influencing the whole decision-making process.

3.3 Protection Motivation Theory in Cybersecurity Research

PMT has gained considerable progress and is now widely acknowledged in the fields of information security and cybersecurity research (Danylak, Lins, & Sunyaev, 2024; Kiran et al., 2025; Tsai et al., 2016). It is one of the most pertinent and resilient theoretical frameworks for elucidating the intentions of individuals to engage in cybersecurity protective behaviors. PMT has been consistently designated as the most frequently applied model in the study of user security behavior over the past two decades in behavioral cybersecurity. This is primarily since PMT's constructs threat appraisal and resilience appraisal are in close alignment with critical cybersecurity concepts (Kiran et al., 2025).

Researchers have applied PMT in diverse contexts, including studies on users’ adoption of security measures, their behavior when exposed to different types of threats, and employees’ perceptions of cybersecurity risks and their coping mechanisms (Li et al., 2019). Additionally, Achuthan et al., 2025 indicated that PMT has been utilized to describe how humans cognitively perceive threats and decide on compliance or non-compliance with security protocols. It has also aided evaluations of public responses to cybersecurity threats and helped identify motivational drivers for preventive activities against technology-based threats (Phan et al., 2025).

Li et al. (2019) discovered that the PMT fundamental aspects of threat assessment and coping appraisal can explain cybersecurity-related behaviors. Across contexts, perceived threat intensity, self-efficacy, and reaction efficacy influence an individual's intention to take preventative security measures. The coping appraisal dimensions particularly self-efficacy and response efficacy, are often found to predict cybersecurity behavior. However, the predictive potential of each PMT variable varies across research and contexts, so the significance of various constructs may rely on threat type, population, or situational factors.

3.4 Strength and Advantages of Protection Motivation Theory

PMT has increasingly gained attention and has been widely applied to information security and cybersecurity research. It is considered one of the most relevant theories for explaining an individual's intention to engage in cybersecurity protective actions. Behavioral cybersecurity research over almost two decades has repeatedly identified PMT as the most frequently used theory in understanding cybersecurity behaviors.

PMT is widely regarded as one of the most robust theoretical frameworks for predicting individuals' intentions to engage in defensive or protective behaviors in the face of perceived threats. Originally developed to explain how persuasive messages, particularly fear appeal, influence behavior, PMT has since evolved into a broader social cognitive model that incorporates both threat appraisal and coping appraisal to understand decision-making in risk contexts (Farooq et al., 2019; Floyd, Prentice-Dunn, & Rogers, 2000; Sulaiman et al., 2023; Tsai et al., 2016).

Empirical evidence supports the theory's validity, with studies consistently reporting statistically significant effect sizes for PMT constructs such as perceived severity, vulnerability, self-efficacy, and response efficacy—aligning with theoretical predictions. This suggests that variations in protective behaviors are strongly associated with these psychosocial factors (Floyd, Prentice-Dunn, & Rogers, 2000). A recent meta-analytic study further confirmed the robustness of PMT across diverse threat contexts (Marikyan & Papagiannidis, 2023).

The framework is considered both comprehensive and intuitive, offering a rational model for weighing the costs and benefits of adopting protective behaviors to mitigate known risks (Fisher, 2024). Its ability to explain the complex cognitive processes underlying protective behavior makes it particularly useful in disciplines such as public health and information security, where it has been extensively used to assess user responses to various security threats (Farooq et al., 2019; Jansen & van Schaik, 2018).

A noteworthy change of PMT was the inclusion of self-efficacy as a fourth cognitive component (Maddux & Rogers, 1983), therefore more closely matching Bandura's Self-Efficacy Theory. Since self-efficacy has been frequently found to be a key determinant of behavioral intention in cybersecurity and health-related settings (Marikyan & Papagiannidis, 2023), this adjustment improved the prediction ability of the model.

While the original PMT model proposed a multiplicative interaction among cognitive variables, the revised version permits researchers to assess the individual contributions of each factor, thereby simplifying empirical testing and improving methodological flexibility (Marikyan & Papagiannidis, 2023; Schneider & Rahman, 2021). Although PMT was initially centered on fear-based appeals, scholars have argued that its components can be extended to broader attitude-change strategies, including those that emphasize positive outcomes with minimal theoretical adjustments (Maddux & Rogers, 1983).

In sum, PMT offers a well-structured and empirically supported approach for examining protective behaviors, especially in cybersecurity research, where understanding user behavior in response to threats is crucial (Fisher, 2024; Jansen & van Schaik, 2018).

3.5 Limitation and Critiques of Protection Motivation Theory

Despite its popularity, PMT has theoretical and methodological flaws. One criticism of PMT is that it does not account for all environmental, cognitive, and moderating elements that affect protective motivation (Marikyan & Papagiannidis, 2023). Although strong evidence that social norms can influence individual behaviour, particularly in group or organisational settings (Almansoori et al., 2023), it omits them. Further research suggests that PMT may not fully explain behavioral intentions in highly specialized situations, necessitating context-dependent variables (Marikyan & Papagiannidis, 2023).

PMT’s core nomology—often simplified in empirical studies—frequently excludes constructs such as maladaptive rewards and the role of fear, both of which were part of the original theoretical formulation. This simplification may reduce the explanatory power of the model in certain applications (Marikyan & Papagiannidis, 2023). Another limitation lies in PMT’s implicit assumption that cognitive processes are invariant across individuals, overlooking the potential moderating roles of personality traits and psychological characteristics. This limits its ability to account for individual differences in decision-making.

Empirical results have also been inconsistent, especially regarding the predictive power of perceived vulnerability and perceived severity (Farooq et al., 2019; Song, Lee, & Roh, 2024). In some cases, threat appraisal constructs have shown insignificant or even negative relationships with behavioral intention. Similarly, response efficacy and response cost have not always emerged as significant predictors. While PMT performs reasonably well in predicting security behaviors, its predictive accuracy suggesting the need to incorporate additional constructs or frameworks to enhance its utility (Marikyan & Papagiannidis, 2023).

Finally, there is a noted lack of intervention studies that test PMT-based mechanisms of change, which hampers efforts to translate theoretical insights into practical behavior-change strategies (Balla & Hagger, 2025). Variation in the effectiveness of PMT components across different cybersecurity contexts may be attributed to contextual factors and methodological differences, further reinforcing the need for contextual tailoring and complementary theoretical perspectives.

4.0 ARTIFICIAL INTELLIGENCE (AI) IN CYBERSECURITY

Artificial Intelligence (AI) is rapidly transforming the landscape of cybersecurity, emerging as a critical tool to combat increasingly sophisticated and evolving cybersecurity threats. Traditional security measures, often based on static rule sets and signature-based detection, have proven inadequate against modern attack vectors. In contrast, AI introduces a paradigm shift by enabling proactive, adaptive, and intelligent defense mechanisms (Akhtar & Rawol, 2024). Nevertheless the reality that AI can be used for positive and negative purposes has caused worry, since criminals can use it as a weapon just as well as good people can use it to protect themselves. (Wilson, 2023; Sivashanmugam & Tan, 2024).

AI significantly enhances cybersecurity by offering a range of advanced applications, including threat detection, anomaly identification, intrusion prevention, malware analysis, phishing and fraud detection, and authentication reinforcement (Aldhamer, 2023). Moreover, AI can automate routine cybersecurity tasks and leverage predictive analytics to anticipate vulnerabilities and emerging attack patterns (Akhtar & Rawol, 2024). These capabilities not only increase efficiency but also allow for real-time and preemptive responses to threats.

Despite its promise, the deployment of AI in cybersecurity is accompanied by several critical challenges and risks. A major concern is the emergence of adversarial AI, where cybercriminals exploit vulnerabilities in AI models to carry out sophisticated, targeted attacks. Additionally, AI system development and operation demand computational resources that may not be widely available (Wilson, 2023; Sivashanmugam & Tan, 2024). Research has found that inconsistent measurement, variability, and overreliance on self-reported data rather than empirical behavioral evidence can affect the reliability and validity of AI research and cybersecurity applications.

Ethical issues also complicate the integration of AI in this field. These include the potential for algorithmic bias, data privacy concerns, and the necessity of human control to prevent misuse. As such, the credibility of AI in cybersecurity remains a double-two-edged sword, balancing promise with threat (Ogundairo & Broklyn, 2024).

Looking ahead, the future of cybersecurity will be increasingly defined by the advancement of AI technologies (Aldhamer, 2023). Acknowledging current limitations is essential for guiding the development of more intelligent, adaptable, and resilient solutions. Future efforts should focus on refining real-time detection mechanisms, integrating emerging AI paradigms, and enhancing the robustness of AI systems to respond to dynamic and evolving threat landscapes. A holistic cybersecurity framework must combine intelligent AI tools with human expertise and intuition, creating a synergistic defense system capable of addressing the multifaceted nature of cyber risks (Akhtar & Rawol, 2024).

5.0 SOCIAL ENGINEERING THREATS IN THE AGE OF AI

Social engineering is commonly regarded as one of the most harmful and persistent dangers in the cybersecurity landscape (Asker et al., 2024; Faotu et al., 2024). It includes persuading individuals into revealing sensitive, confidential, or personal information, using human vulnerabilities rather than technical ones. Despite breakthroughs in security hardware and software, cybercriminals continue to evade technical barriers by targeting the human factor, which remains a significant vulnerability in most systems (Manyam, 2022; Asker et al., 2024).

The rise of AI has significantly transformed the nature and effectiveness of social engineering attacks. While AI offers remarkable capabilities in defense such as threat detection, behavioral analytics, and automation it simultaneously equips cybercriminals with enhanced tools to execute more targeted and convincing attacks. This shows that AI is a “two-edged sword” when it comes to defense (Wilson, 2023; Manyam, 2022; Pujari & Hussain, 2024).

In several factors, AI makes social engineering more apparent and broader. Using natural language processing (NLP) and machine learning (ML), scammers can produce phony emails that look like they originated from trustworthy contacts or closely mirror the writing style of their targets (Bauskar et al., 2024; Vulpe et al., 2024). Publicly available data, especially from social media, is collected and examined to develop attacks that are very personalized to each person, which escalates the risk of lying. Also, AI-driven bots can now talk to people like humans through chat interfaces, making it harder to distinguish the difference between real and fraudulent interaction.

More concerning deepfakes and voice cloning can create convincing multimedia material that fools even vigilant consumers. Financial frauds use AI-generated voices to imitate executives, and chatbot-equipped phishing websites mimic customer assistance conversations (Manyam, 2022; Wilson, 2023). These capabilities simplify and lower the cost of large-scale, effective attacks. AI improves attack precision, speed, and scalability. AI systems can expand spear phishing and CEO fraud, altering content and tone based on real-time feedback. Cybercriminals can now easily launch credible and context-aware phishing attacks thanks to large language models (LLMs) (Faotu, Asheshemi, & Jeremiah, 2024; Wilson, 2023).

Defending against AI-driven social engineering poses unique challenges. Traditional cybersecurity methods that focus on technical vulnerabilities fall short when confronting attacks rooted in human psychology. Detection remains difficult, as there is no universal model capable of dynamically identifying evolving social engineering patterns (Faotu et al., 2024). Although research has explored automated detection through NLP and ML, the success of these systems remains limited due to the ever-changing tactics of cybercriminals.

To counter these threats, the integration of AI in cybersecurity defense systems is not only beneficial but essential. AI can analyze behavioral patterns, predict threats, and recommend proactive measures (Aldhamer, 2023). ML algorithms serve as crucial countermeasures to reduce the effectiveness of socially engineered attacks by detecting subtle anomalies and inconsistencies (Chouraik, El-Founir, & Taibi, 2024). Furthermore, AI-driven personalized training and awareness programs can help close gaps in digital literacy, tailoring interventions based on an individual’s learning pace and susceptibility to manipulation (Faotu et al., 2024).

Nonetheless, reliance on AI alone is insufficient. Building a robust human firewall requires a synergistic approach—combining advanced technical defenses with continuous user education, awareness, and behavioral reinforcement (Ogundairo & Broklyn, 2024; Al Siam, Alazab, Awajan, & Faruqui, 2025). Moving forward, strategic AI integration must be guided by ethical considerations, transparency, and a commitment to preserving user trust and data privacy.

6.0 CONCLUSION

Despite its importance in the face of emerging threats, cybersecurity behavior research, particularly on the human aspect and AI, has some inherent limitations. The frequent use of cross-sectional data, which collects information at a single point in time, limits the capacity to establish causal linkages or account for changes over time. Without experimental controls, external variables may affect study results, making causation harder to prove. Self-reported behavioral data is often used, which can be biased due to social desirability.

Previous research is limited by theoretical frameworks and scope. Studies often focus on PMT without integrating other relevant behavioral theories or considering the full nomology of constructions, potentially overlooking social norms, moral obligations, or personality traits that could provide additional insights. Maladaptive response rewards and response costs, two crucial PMT variables, have not been widely explored, and threat data is inconsistent. Technological improvements and cybersecurity threats grow regularly, therefore findings can become outdated. Cybersecurity behavior and AI research should focus on three critical areas to overcome these limits. Experimental and continuous research approaches would help understand causal pathways and how awareness and behavior evolve over time. Researchers can also apply non-self-report behavior indicators to eliminate self-report bias and better appreciate the intention-behavior gap. Expanding the demographic coverage beyond students and professionals in developed nations is essential to improve external validity and understand how cultural contexts and socioeconomic backgrounds affect cybersecurity perceptions and behaviors. Mixed techniques, integrating quantitative and qualitative data, may help explain the complex aspects affecting cybersecurity awareness and behavior.

Future research should also include social, psychological, and technical factors, moderators, and mediators like fear in behavioral theories. Training programs and interventions must also be evaluated. In the future, researchers should work on making AI models better at adapting to new dangers, being clear, being able to understand, being ethically sound (including privacy and bias), and stopping attacks from other AI models. Predictive modeling can benefit from using machine learning algorithms to evaluate vast volumes of data and detect cybersecurity behavior patterns. Finally, to stay current with cybersecurity threats and future technologies like quantum computing, security methods and understanding must be researched continuously.

REFERENCES

Achuthan, K., Khobragade, S., & Kowalski, R. (2025). Public sentiment and engagement on cybersecurity: Insights from Reddit discussions. *Computers in Human Behavior Reports*, *17*, 100573.

Akhtar, Z. B., & Rawol, A. T. (2024). Enhancing cybersecurity through artificial intelligence (AI)-powered security mechanisms. *IT Journal Research and Development*, *9*(1), 50–67.

Alghazo, S. H. A., Humaidi, N., & Abdulla, N. B. (2025). Utilising Manager's Competency, Employee's Awareness and Motivation for Promoting Cybersecurity Protective Behaviour. *The Electronic Journal of Knowledge Management*, *23*(2), 14-40.

Al Siam, A., Alazab, M., Awajan, A., & Faruqui, N. (2025). A comprehensive review of AI's current impact and future prospects in cybersecurity. *IEEE Access*, *13*, 14029–14050

Al-Harthy, I. M., Abdul Rahim, F., Ali, N., & Singun Jr., A. P. (2020). Dimensions of protection behaviors: A systematic literature review. *Journal of Theoretical and Applied Information Technology*, *98*(17), 3668–3697.

Al-Hashem, N., & Saidi, A. (2023). *The psychological aspect of cybersecurity: Understanding cyber threat perception and decision-making*. International Journal of Applied Machine Learning and Computational Intelligence, 13(8), 11–22.

Aldhamer, M. (2023). The impact of artificial intelligence on the future of cybersecurity. *Middle East College Student Journal*, 1(1), 1–15.

Almansoori, A., Al-Emran, M., & Shaalan, K. (2023). Exploring the Frontiers of Cybersecurity Behavior: A Systematic Review of Studies and Theories. *Applied Sciences*, *13*(9), 5700.

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), 100463.

Asker, N. S., et al. (2024). A review of social engineering attack detection based on machine learning techniques. *Nanotechnology Perceptions*, 20(S7), 675–686.

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.

Bauskar, S. R., Madhavaram, C. R., Galla, E. P., Sunkara, J. R., & Gollangi, H. K. (2024). AI-driven phishing email detection: Leveraging big data analytics for enhanced cybersecurity. *Library Progress International*, *44*(3), 7211–7224.

Chouraik, C., El-Founir, R., & Taibi, K. (2024). The impact of AI on cybersecurity: A new paradigm for threat management. *African Journal of Management, Engineering and Technology*, *2*(2), 92–99.

Danylak, P., Lins, S., & Sunyaev, A. (2024). The role of employees' threat appraisal in security certification compliance: Insights from a protection motivation approach. In *Proceedings of the 57th Hawaii International Conference on System Sciences (HICSS)*. University of Hawai'i at Mānoa.

Debb, S. M., & McClellan, M. K. (2021). Perceived vulnerability as a determinant of increased risk for cybersecurity risk behavior. *Cyberpsychology, Behavior, and Social Networking*, *24*(9), 605–611.

Dong, H., & Kotenko, I. (2025). Cybersecurity in the AI era: Analyzing the impact of machine learning on intrusion detection. *Knowledge and Information Systems*, *67*(5), 3915–3966.

Faotu, H., Asheshemi, O. N., & Jeremiah, T. E. (2024). Human vulnerabilities in cybersecurity: Analyzing social engineering attacks and AI-driven machine learning countermeasures. *Journal of Science and Technology*, *30*(1), 72–84.

Farooq, A., Ndiege, J. R. A., & Isoaho, J. (2019, September). *Factors Affecting Security Behavior of Kenyan Students: An Integration of Protection Motivation Theory and Theory of Planned Behavior*. In 2019 IEEE AFRICON (pp. 1-6)

Fisher, T. S. (2024). *Gone smishing: An integration of decision-making and protection motivation in the context of identity theft victimization* (Doctoral dissertation). University of South Florida.

Floyd, D. L., Prentice-Dunn, S., & Rogers, R. W. (2000). A meta-analysis of research on protection motivation theory. *Journal of Applied Social Psychology*, *30*(2), 407–429.

Jabali, A., & Baher, N. (2024). *Exploring the impact of cybersecurity knowledge and awareness on behavioral choices for protection among university students in Sweden* (Master's thesis, Luleå University of Technology). DiVA Portal.

Jansen, J., & van Schaik, P. (2016). Understanding precautionary online behavioural intentions: A comparison of three models. In *Proceedings of the Tenth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2016)* (pp. 1–11). Frankfurt, Germany.

Jansen, J., & van Schaik, P. (2018). Testing a model of precautionary online behaviour: The case of online banking. *Computers in Human Behavior, 87,* 371–383

Khadka, K., & Ullah, A. B. (2025). Human factors in cybersecurity: an interdisciplinary review and framework proposal. *International Journal of Information Security*, *24*(1), 119.

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,* 104204.

Li, L., He, W., Xu, L., Ash, I., Anwar, M., & Yuan, X. (2019). Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior. *International Journal of Information Management, 45,* 13–24.

Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology, 19*(5), 469–479.

Manyam, S. (2022). *Artificial intelligence's impact on social engineering attacks* (Master's capstone project, Governors State University). OPUS Open Portal to University Scholarship.

Marikyan, D., & Papagiannidis, S. (2023). Protection Motivation Theory: A review. In S. Papagiannidis (Ed.), *TheoryHub Book*.

Ogundairo, O., & Broklyn, P. (2024). *AI-driven phishing detection systems*. EasyChair Preprint No. 14338.

Phan, B. T., Do, P. H., & Le, D. Q. (2025). The Impact Of Digital Literacy On Personal Information Security: Evidence From Vietnam. In D. N. Van et al. (Eds.), *Proceedings of the International Conference on Emerging Challenges: Sustainable Strategies in the Data-driven Economy (ICECH 2024), Advances in Economics, Business and Management Research 320*.

Pujari, S. R., & Hussain, M. A. (2024). Human factor in cybersecurity: Behavioral insights into phishing and social engineering attacks. *Nanotechnology Perceptions, 20*(S15), 630–642.

Rogers, R. W. (1983). Cognitive and physiological processes in fear appeals and attitude change: A revised theory of protection motivation. In J. Cacioppo & R. Petty (Eds.), *Social psychophysiology: A sourcebook* (pp. 153–177). Guilford Press.

cybersecurity: a comprehensive review of AI-driven detection techniques. *Journal of Big Data*, *11*(1), 105.

Schneider, M., & Rahman, S. (2021). Protection motivation theory factors that influence undergraduates to adopt smartphone security measures. *IT in Industry, 9*(1), 1–12.

Sivashanmugam, L., & Tan, J.-E. (2024). *Cybersecurity risks of AI*. Khazanah Research Institute.

Sulaiman, N. S., Fauzi, M. A., Hussain, S., & Wider, W. (2022). Cybersecurity Behavior among Government Employees: The Role of Protection Motivation Theory and Responsibility in Mitigating Cyberattacks. *Information*, *13*(9), 413.

Sulaiman, N. S., Aziz, N. S., Nasir, A., & Yacob, A. (2023). Cyber security awareness model (among children) using Protection Motivation Theory: A review. *International Journal of Business and Technology Management, 5*(1), 74–85.

Tsai, H. Y. S., Jiang, M., Alhabash, S., LaRose, R., Rifon, N. J., & Cotten, S. R. (2016). Understanding online safety behaviors: A protection motivation theory perspective. *Computers & Security*, *59*, 138–150.

Vafaei-Zadeh, A., Nikbin, D., Teoh, K. Y., & Hanifah, H. (2025). Cybersecurity awareness and fear of cyberattacks among online banking users in Malaysia. *International Journal of Bank Marketing*, *43*(3), 476–505.

Verma, M. K., & Kushwaha, S. S. (2025). Exploring the latent interrelationship of cybercrime awareness and personality traits: A review. *Asian Journal of Education and Social Studies, 51*(5), 347–354.

Vulpe, S.-N., Rughiniș, R., Turcanu, D., & Rosner, D. (2024). AI and cybersecurity: a risk society perspective. *Frontiers in Computer Science*, *6*, 1462250.

Wilson, S. (2023). *Cybersecurity and artificial intelligence: Threats and opportunities*. Contrast Security.