Human Vulnerabilities to Social Engineering Attacks: A Systematic Literature Review for Building a Human Firewall

Muhammad Shofian Tsauri
Sistem Informasi, UIN Syarif Hidayatullah Jakarta
shofian.ts21@mhs.uinjkt.ac.id

Article Info

Article history:

Received 2025-05-01
Revised 2025-06-02
Accepted 2025-07-04

Keyword:

Social Engineering,
Information Security,
Human Vulnerability,
Systematic Literature Review,
Human Firewall.

ABSTRACT

Social engineering attacks exploit human psychology to deceive individuals into compromising information security, making the human element a critical vulnerability in cybersecurity systems. This study aims to identify and analyze patterns of human susceptibility in social engineering through a systematic literature review (SLR). Guided by the PRISMA 2020 protocol, a total of 865 articles were initially retrieved from databases such as Scopus, IEEE Xplore, ResearchGate, and Google Scholar. After applying strict inclusion and exclusion criteria, 39 peer-reviewed articles published between 2020 and 2024 were selected for thematic synthesis. The results reveal recurring human vulnerability factors including low security awareness, emotional manipulation (e.g., fear, urgency), overtrust in authority, and lack of behavioral control. These vulnerabilities manifest in predictable victim profiles and behavioral patterns, which are often exploited through phishing, pretexting, and other deception-based tactics. Furthermore, the review highlights the limitations of current mitigation strategies that focus solely on technical solutions without integrating human behavior models. The findings serve as a conceptual foundation for building a "human firewall," emphasizing awareness, vigilance, and behavioral training as integral components of social engineering defense. This study also lays the groundwork for the development of a human-centric detection model in future research, particularly in the context of mobile banking.

I. INTRODUCTION

The rapid advancement of information technology has brought remarkable convenience to various aspects of human life, ranging from banking and business activities to everyday communication. However, alongside this progress, information security threats have also increased, one of which is in the form of social engineering attacks. Social engineering is an attack method that exploits human psychological vulnerabilities to gain unauthorized access to systems or sensitive information. [1].

According to the 2022 Verizon Data Breach Investigations Report (DBIR), approximately 82% of security breaches involved human factors, including social engineering, human error, and misuse of access privileges. [2]. In addition, a 2023 report by KnowBe4 revealed that phishing attacks—one of the primary forms of social engineering—increased by 61% compared to the previous year, resulting in global financial losses amounting to billions of dollars. [3].

This condition indicates that the human factor remains the most vulnerable element in modern information security systems. Amid advances in security technologies such as encryption and multi-factor authentication, human-based attacks continue to demonstrate high effectiveness. This highlights the need for greater emphasis on behavior-based security approaches and user awareness.

Although numerous studies on social engineering have been conducted, most of them still focus on the technical aspects of protection or merely describe the types of attacks. There remains a lack of systematic reviews that specifically examine how individuals fall victim to social engineering attacks, as well as the psychosocial factors that influence their vulnerability.

Based on this background, this study aims to conduct a Systematic Literature Review (SLR) on the factors that contribute to individuals' vulnerability to social engineering attacks. Through this approach, it is expected to produce a structured mapping of human vulnerability patterns, which can later serve as a foundation for developing a human firewall-based defense model in future research.

In line with this objective, the study is designed to address the following two research questions:

- RQ1: How do individuals fall victim to social engineering attacks in the context of information security?
- RQ2: What factors increase individuals' vulnerability to social engineering attacks?

Thus, this study is expected to make a scientific contribution by enriching the discourse on human vulnerabilities in the field of information security, as well as serving as a reference for developing human-centered mitigation strategies to address social engineering threats in the digital era.

II. METHOD

This research uses the Systematic Literature Review (SLR) method to examine the factors that cause individuals to fall victim to social engineering attacks. This approach was chosen to obtain a systematic and comprehensive synthesis of previous research in the field of human-centered information security. The flowchart is presented as follows:

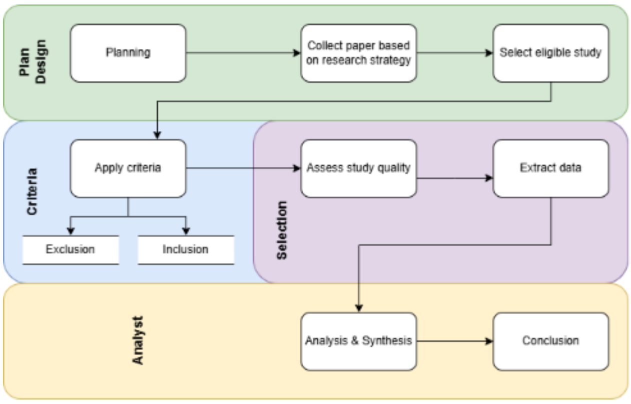

Figure 1. Diagram Alur SLR

A. Plan Design

The design of this study focuses on the collection, selection, analysis, and synthesis of data from various academic literature related to social engineering. The SLR method enables researchers to identify key patterns in human vulnerability based on published empirical findings.

The search strategy was conducted across several reputable academic databases, including:

- Google Scholar
- IEEE Xplore
- Scopus
- ResearchGate

To ensure relevance and comprehensiveness, search queries were built using Boolean operators and keyword combinations that reflect the core focus of the study. The main keywords used include:

- "Social Engineering Attack"
- "Human Vulnerability in Information Security"
- "Victim Psychology Social Engineering"
- "Social Engineering Prevention" "User Behavior Security Awareness"
- "Phishing Sesceptibility"
- "Human Error in Cybersecurity"

Boolean operators (AND/OR) were used to refine result, for exampe:

```
("social engineering" OR phishing OR smishing)
AND
("human vulnerability" OR "user behavior" OR
"security awareness")
AND
("information security" OR cybersecurity)
```

B. Criteria

The literature selection process was conducted based on clearly defined inclusion and exclusion criteria to ensure that the articles selected were relevant, reliable, and of high academic quality. These criteria helped filter the most appropriate studies to answer the research questions and ensure alignment with the systematic review protocol

1) *Inclusion Criteria*: Studies were included in the final review if they met the following criteria:

- Type of Document: Empirical research in the form of peer-reviewed journal articles or conference proceedings.
- Topic Relevance: Studies that specifically focus on human vulnerabilities in the context of social engineering attack, such as phishing, smishing, baiting, pretexting, or impersonation.
- Content Scope: Articles that examine vehavioral, psychological, cognitive, or social factors contributing to human susceptibility to social engineering.
- Publication Year: Articles published between 2020 and 2024, to ensure the relevance of the data with current threat landscapes and user behavior trends.

2) *Exclusion Criteria*: Studies were excluded from the review based on the following conditions:

- Non-Empirical Nature: Literature that does not provide empirical data, such as opinion pieces, blog post, white papers, editorials, or purely theoretical essays.
- Lack of Peer Review: Articles that have not undergone academic peer review, including preprints or informal publications.

- Irrelefant Focus: Studies that focus solely on technical aspects of information security (e.g., firewalls, encryption, network protocols) without addressing the human element.
- Redudancy: Duplicate studies or publications with overlapping datasets from the same research.

C. Selection

To ensure methodological rigor and transparency, the study selection process was conducted following the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The selection process was divided into four main phases: Identification, Screening, Eligibility, and Inclusion.

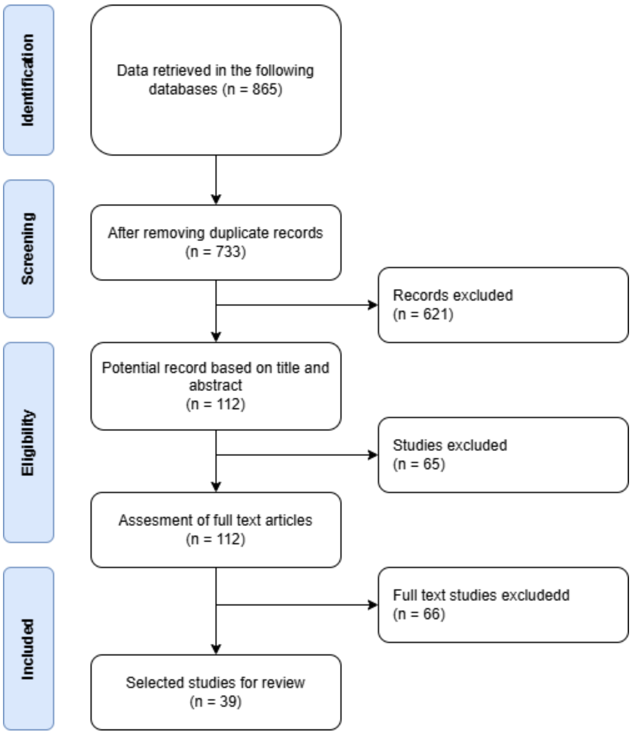
Identification
Screening
Eligibility
Included
Data retrieved in the following databases (n = 865)
After removing duplicate records (n = 733)
Records excluded (n = 621)
Potential record based on title and abstract (n = 112)
Studies excluded (n = 65)
Assesment of full text articles (n = 112)
Full text studies excludedd (n = 66)
Selected studies for review (n = 39)

Figure 2. Four Flow Selection of Prisma Guidelines

The following steps were taken:

1) *Identification*: An initial search across four major academic databases (Google Scholar, IEEE Xplore, Scopus, and ResearchGate) produced a total of 865 articles. After removing 132 duplicates, 733 unique articles remained for further screening.

2) *Screening*: Titles and abstracts of the remaining articles were reviewed for relevance to the research questions. 621 articles were excluded at this stage for reasons such as irrelevant scope (technical focus only), language issues, or lack of human factor discussion. This left 112 articles for full-text eligibility assessment.

3) *Eligibility*: A detailed review of the full texts of these 112 was conducted using the inclusion and exclusion criteria. 65 articles were excluded due to one or more of the following reasons:

- No empirical findings
- Conceptual or theoretical only
- Unavailable full text
- Poor methodological quality

4) *Inclusuion*: The final set included 39 studies that met all criteria and were considered suitable for data extraction and thematic synthesis.

The article selection process consists of two main steps. First, screening based on titles and abstracts to determine initial relevance. Second, a full-text examination of the articles to ensure alignment with the predefined inclusion and exclusion criteria. The publication date range is limited to 2020 through 2024 to ensure relevance to current conditions.

D. Analyst

The data obtained from the selected literature is extracted and categorized based on the identified vulnerability factors. The analysis is conducted qualitatively using a narrative synthesis method to group common patterns related to how individuals become victims of social engineering attacks.

TABLE I
LIST OF PAPERS

Citation	Title	Quality
Journal Q1 (Top Tier)		
[4]	Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods	Q1 (Scopus/WoS)
[5]	Social engineering attacks prevention: A systematic literature review	Q1 (Scopus/WoS)
[6]	Spear-phishing susceptibility stemming from personality traits	Q1 (Scopus/WoS)
[7]	Performing social engineering: A qualitative study of information security deceptions. Computers in Human Behavior	Q1 (Scopus/WoS)
[8]	2020). Predicting individuals' vulnerability to social engineering in social networks	Q1 (Springer)
Journal Q2 (Mid Tier)		
[9]	Study on the psychology of social engineering-based cyberattacks and existing countermeasures	Q2 (Scopus)
[10]	Overview of social engineering attacks on social networks	Q2 (Proceeding Scopus)
[11]	Social engineering attacks: Recent advances and challenges	Q2 (Springer LNCS)
[12]	Systematic review on social engineering: Hacking by manipulating humans	Q2 (Scopus)

[13]	Reviewing cyber security social engineering training and awareness programs-pitfalls and ongoing issues	Q2
Journal Q3-Q4		
[14]	Coronavirus social engineering attacks: Issues and recommendations	Q4
[15]	An advanced taxonomy for social engineering attacks	Q4
[16]	Social engineering cyber threats	Q4
[17]	Analisis indikator utama dalam information security: Personality threat terhadap phishing attack	Q4
[18]	A comprehensive survey of phishing attacks and defences: Human factors, training and the role of emotions	Q4
[19]	The psychology of social engineering	Q4
[20]	A comprehensive survey on social engineering-based attacks on social networks	Q4
[21]	Personal data vulnerability in the digital era: Study of modus operandi and mechanisms to prevent phishing crimes	Q4
[22]	Phishing attacks in social engineering: A review	Q3
Conference and Proceeding		
[23]	Mitigating social engineering attack: A focus on the weak human link	Conference IEEE (Q3)
[24]	Analysis of social engineering attacks using exploit kits	Proceeding Springer (Q3).
[25]	Human susceptibility to phishing attacks based on personality traits: The role of neuroticism	Conference IEEE (Q3).
[26]	Impact of social engineering attacks: A literature review	Q3 (Proceeding Springer).
[27]	Predicting personal susceptibility to phishing	Q3 (Proceeding Springer, terindeks Scopus).
[28]	Social engineering attacks: A reconnaissance synthesis analysis	Conference IEEE (Reputable).
[29]	Social engineering: The looming threat	Conference IEEE (Reputable).
[30]	The main social engineering techniques aimed at hacking information systems	Conference IEEE (Reputable).
[31]	Human emotion factor based classification and defense against social engineering attacks	Conference IEEE (Reputable).
[32]	Impact analysis and performance model of social engineering techniques	Conference IEEE (Reputable).
[33]	Comprehensive assessment of reverse social engineering to understand social engineering attacks	Conference IEEE (Reputable).
[34]	Internet-based social engineering attacks, defenses and psychology: A survey	Preprint
Other Researchers		
[35]	Psychological aspects of the organization's information security in the context of socio-engineering attacks	Q4 (Psychological)
[36]	Improving social engineering resilience in enterprises	Q3 (Jurnal baru, belum terindeks).
[37]	Non-technical cyber-attacks and international cybersecurity: The case of social engineering	Q4 (Non-Technical).
[38]	Social engineering attacks and countermeasures	Q3 Book chapter (IGI Global)
[39]	Innovations of phishing defense: The mechanism, measurement and defense strategies	Q3 (Low Impact)
[40]	Mitigating social engineering attacks on the elderly: Personalized countermeasures to enhance cyber situational awareness	Q4 Conference non-teknikal
[41]	A mathematical model for risk assessment of social engineering attacks	Q3 (Low Impact)
[42]	The identification of a model victim for social engineering: A qualitative analysis	Q2 (Fokus psikologi)

III. Results And Discussion

A. Summary of Literature Study

This study examines 39 scholarly articles related to human vulnerabilities to social engineering attacks. These articles were obtained through a systematic search across reputable academic databases such as Google Scholar, IEEE Xplore, Scopus, and ResearchGate, with a publication date range from 2020 to 2024.

The selection process was conducted based on the predefined inclusion and exclusion criteria. From this process, relevant and eligible articles for further analysis were obtained.

A variety of research methods were used in the articles, which can be summarized as follows.

TABLE II
SUMMARY OF METHODS IN LITERATURE PAPERS

Method	Paper
Phishing Analysis and SE Techniques	9
Surveys and Questionnaires	5
Prediction Models	4
Case Studies	5
Frameworks	6
Experiments	4
Phishing Campaign Analysis	7
General Literature Reviews	5

And it provides contributions as well as suggestions for future research as follows:

1) *Contribution*: A total of 35 out of 39 articles provide explicit contributions in the form of model development, identification of vulnerability factors, or proposed frameworks for mitigating social engineering.

2) *Suggestions for Future Research*: A total of 18 out of 39 articles suggest further research, such as empirical experiments to deepen the understanding of human factors and the development of human-centric approaches to enhance information security awareness.

The following are the first 15 out of 39 papers reviewed by the author (for complete data, please contact the author via email):

TABLE III
15/39 PAPERS USED IN THE SLR

	Citation	How Individuals Can Become Victims	Methods	Suggestions for Future Research
1	[4]	Individuals become victims due to various forms of social engineering attacks.	13 types of social engineering attacks.	Further empirical studies on SE attacks are needed.
2	[5]	Individuals become victims due to ignorance or negligence towards SE attacks.	Systematic Literature Review.	Exploring work guidelines for humans as security sensors.
3	[10]	Victims fall prey to psychological manipulation and exploitation.	Affects the target, psychological exploitation.	Design systematic steps for responding to SE attacks.
4	[9]	This is mentioned generally without technical details.	Phishing, pretexting, baiting, tailgating.	Not mentioned.
5	[8]	It does not directly discuss how individuals become victims.	User vulnerability prediction models.	Not mentioned.
6	[14]	Fear during the COVID-19 pandemic was exploited by SE perpetrators.	Simulation training, interactive, and online.	Analysis of detection techniques and implementation of preventive measures.
7	[11]	Manipulation techniques and human-computer interaction exploit psychological vulnerabilities.	Deception, psychological manipulation.	Not mentioned.
8	[15]	Exploitation through VoIP, SMSishing, and manipulation techniques.	Advanced SE attack taxonomy.	Vulnerability factors and the prevalence of preventive measures.
9	[29]	Curiosity and trust are exploited through various vectors.	Attacks through observation, dumpster diving, and other simple techniques.	Not mentioned.
10	[28]	Phishing, pretexting, and baiting techniques exploit trust and fear.	Reconnaissance Synthesis Analysis.	Not mentioned.
11	[12]	Manipulative techniques such as phishing and pretexting, psychological factors of trust and urgency.	Shoulder surfing and dumpster diving.	Education and incident response strategies for SE.
12	[30]	Phishing techniques (vishing, smishing, spear phishing, etc.) exploit curiosity and fear.	Various forms of phishing.	Not mentioned.
13	[7]	Psychological exploitation through phishing,	Grounded theory from interviews with 37 SE perpetrators.	Effectiveness of SE techniques and user training methods.

		baiting, and pretexting.		
14	[6]	Personality traits (neuroticism, agreeableness) affect SE vulnerability.	Spear-phishing campaigns in software companies.	Email phishing tailored to personality types.
15	[26]	Lack of training and naivety create gaps for SE.	Systematic literature review.	Not mentioned.

B. Patterns of Individuals Becoming Victims

Based on the review of 39 articles, several common patterns were identified that explain how individuals can become victims of social engineering attacks. These patterns are closely related to psychological characteristics, social conditions, and users' digital interactions. Social engineering attackers consistently exploit these gaps to build trust and lead victims to take certain actions, such as revealing confidential information or accessing malicious links. The summary is as follows:

TABLE IV
PATTERNS OF INDIVIDUALS FOUND IN THE LITERATURE

General Pattern	Desciption	Citation
Emotions (fear, anxiety, panic)	Encouraging the victim to react without thinking critically	[10], [17], [19], [35], [18], [20], [42]
Abuse of authority	Perpetrators posing as official institutions or important individuals	[4], [6], [22], [37], [24]
Lack of security awareness	Victims not understanding the risks or characteristics of social engineering	[8], [5], [40], [33], [36], [21], [38], [27]
Fake urgency	Tactics that force the victim to act immediately	[26], [23], [34], [41], [24]
Digital and cognitive fatigue	Victims making wrong decisions due to information overload	[39], [37], [25]

1) Utilization of Emotions and Psychology:
Research indicates that many individuals become victims due to being influenced by certain emotional conditions. Emotions such as fear, urgency, and trust are intensively exploited by social engineering perpetrators. For example, [10], [17], and [19] It is noted that a common tactic in phishing is to create the illusion that the victim's account will be blocked if immediate action is not taken.

2) Abuse of Authority and Trust:
Many social engineering perpetrators impersonate authoritative figures (e.g., bank officers, workplace supervisors, or IT technicians). Studies in [4], [6], and [22] It is noted that victims are more likely to comply with instructions when perpetrators claim to be from official institutions or use a professional communication style.

3) Lack of Security Awareness:
The majority of victims are unaware of social engineering threats, especially those without a technological background or who have never received security training. Articles [8], [5], and [40] It is identified that victims often fail to recognize common signs of phishing or online fraud.

4) Urgent Situations:
Several studies, such as [26], [23], and [34] highlight that social engineering perpetrators often create urgent conditions (e.g., "Your account will be blocked in 10 minutes" or "Immediately confirm this OTP code") to prompt victims to act without critical thinking.

5) Information Overload and Fatigue:
Cognitive factors also play a role, where victims tend to make errors when experiencing high workloads or digital fatigue. Studies such as [39] and [25] note that under these conditions, victims are more likely to click on malicious links or share sensitive data.

These patterns indicate that the success of social engineering attacks is highly influenced by psychological and situational human factors. Attackers tend to avoid technical aspects and instead prefer to exploit human nature aspects such as fear, trust, and weak digital habits.

C. Individual Vulnerability Factors

It was found that individual vulnerability to social engineering attacks is influenced by various multidimensional factors. These factors can be classified into three main categories: psychological, social, and digital-behavioral. These three categories are interconnected and form the context in which individuals become more susceptible to manipulation by social engineering perpetrators. The summary is as follows:

TABLE V
VULNERABILITY FACTORS IDENTIFIED IN THE LITERATURE

Chategory	Sub-Factors	Citation
Psychological	Fear, panic, excessive trust, obedience	[4], [10], [6], [17], [22], [23], [19], [34], [20]
Social	Social pressure, cultural norms, low awareness	[8], [11], [26], [40], [31], [22]
Digital-Behavioral	Random clicking, low literacy, password reuse	[5], [8], [11], [13], [7], [39], [32], [16], [23]

1) Psychological Factors:
This factor relates to the mental and emotional state of the individual. Social

engineering perpetrators often exploit the victim's emotions to trigger impulsive responses. Such as (a) Fear and panic [[10], [17]]; (b) Excessive trust in seemingly credible parties [[4], [22]]; (c) Obedience to authority [[6], [23]]; (d) Overconfidence or a sense of security [[19], [34]].

2) Social Factors: Refers to the influence of the social environment and cultural norms that affect an individual's perception and response. Such as (a) Peer or group pressure [[8], [40]]; (b) Social norms of "not questioning authority" [[[[26], [22]]; (c Lack of security education in communities or organizations [[11], [31]].

3) Digital-Behavioral Factors: This factor encompasses an individual's habits, skills, and digital literacy, which affect how they interact with technology. Such as: (a) Lack of digital security awareness [[5], [8], [7]]; (b) Habit of randomly clicking on links [[11], [32]]; (c) Failure to verify the identity of message senders [[13], [23]]; (d) Reusing passwords / not using two-factor authentication [[39], [16]]

Psychological factors often serve as the starting point for attacks, particularly when combined with certain social situations (e.g., when victims are pressured to respond promptly to seemingly official requests). Meanwhile, digital-behavioral factors reveal that poor online habits are a common gateway for attackers. Many studies emphasize that the combination of a lack of awareness and social pressure can create an optimal condition for the success of social engineering attacks.

D. Research Implications

Based on the review findings, it can be concluded that the mechanism through which victims fall into social engineering traps (RQ1) is closely linked to both internal and external vulnerability factors possessed by individuals (RQ2). For example, behaviors such as responding to messages impulsively, as discussed in section B, are often driven by a combination of psychological factors, such as fear outlined in section C, lack of digital literacy, and social pressure.

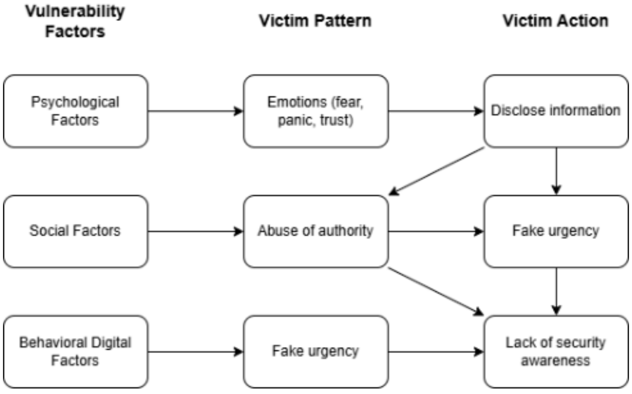

Figure 3. Relationship Diagram Between Factors, Patterns, and Victim Actions.

The results of the exploration conducted by the author revealed three major themes that emerge in the integration of both factors, as follows:

1) *Emotional Exploitation:* Perpetrators exploit the victim's emotional state, such as panic or excessive trust.

2) *Low Security Literacy:* Victims are unaware of the signs of social engineering or do not understand how to protect themselves.

3) *Lack of Social and Organizational Protection:* There is no training system or security culture in the workplace.

The findings of this systematic literature review have significant implications for various stakeholders, particularly in designing behavioral security strategies and enhancing resilience against social engineering attacks. The review identified that individual susceptibility is not merely accidental but is shaped by predictable patterns of behavior, emotional responses, and cognitive vulnerabilities. These insights have both theoretical and practical implications in the broader context of human-centric cybersecurity.

1) *Implications for Users (End-Users or Victims)*: The findings suggest that users remain the weakest link in cybersecurity ecosystems due to limited awareness, overtrust, and emotional triggers such as urgency, authority, and fear:

- Security awareness campaigns must be tailored to cognitive and emotional patterns, not just technical instructions.
- Repetitive and contextual phishing simulation training can help users recognize attack patterns over time.
- Tools such as gamified training modules or microlearning via mobile apps could reinforce secure behavior.

2) *Implications for Institutions and Organizations*: Organizations need to shift their defense strategies from system-centric to user-inclusive models. Based on this review:

- Institutions should integrate behavioral-based security education as part of onboarding and periodic training.
- Policies must be oriented not only on compliance but also on behavior change supported by measurement and feedback systems.
- A culture of security should be cultivated where employees are empowered to question suspicious activity without fear of reprimand.

3) *Implications for System Developers and Security Designer*: Security features must be designed with human limitations in mind. This includes:

- Designing interfaces that reduce cognitive overload, highlight suspicious behavior, and provide just-in-time warnings.

- Incorporating anomaly-based detection that tracks user behavior and flags deviations for verification.
- Allowing for adaptive systems that offer context-sensitive prompts (e.g., “Are you sure this message is from your bank?”) before action is taken.

This synthesis serves as the foundational basis for the development of the Social Engineering Attack Detection Model (SEADM) based on a Human Firewall, which will be discussed in subsequent research. By thoroughly understanding the patterns and vulnerability factors, the development of preventive models can be carried out in a more contextual and data-driven manner.

E. Conceptualizing the Human Firewall

The term *human firewall* has emerged as a conceptual response to the persistent exploitation of human vulnerabilities in social engineering attacks. Unlike technical firewalls that operate through code and network configurations, a human firewall refers to the cognitive, behavioral, and emotional resilience of individuals when confronted with manipulative attack vectors. It represents the capacity of end-users to recognize, question, and resist deceptive tactics commonly used in phishing, pretexting, and other social engineering methods.

Several studies included in this review emphasize the pivotal role of human behavior and awareness in either enabling or mitigating social engineering risks. The recurring themes in the reviewed literature suggest that the human firewall is not a singular attribute but a composite of multiple psychological and educational factors, including:

- Security Awareness: The degree to which users understand threats and their potential impact.
- Behavioral Compliance: Adherence to organizational policies, security protocols, and warning cues.
- Emotional Regulation: The ability to maintain rational judgment under emotional manipulation.
- Cognitive Vigilance: The skill to recognize anomalies, verify identities, and avoid impulsive actions.
- Security Habituation: The result of repeated exposure to simulations or training that reinforces secure behavior.

This synthesis of findings supports the view that a human firewall must be built intentionally, through targeted training, policy design, and behavioral modeling, rather than assumed as an inherent user quality. While most technical controls operate at the system level, the human firewall operates at the decision-making level, influenced by culture, context, and education.

Thus, the outcomes of this SLR provide a conceptual scaffold for building the human firewall. These insights will serve as the foundational basis for developing a human-centric detection framework in the authors’ subsequent research, titled *Social Engineering Attack Detection Model (SEADM) Based on Human Firewall in M-banking Services.*

F. Future Research

While this study provides a comprehensive synthesis of human vulnerabilities exploited in social engineering attacks, it also opens several avenues for further research. One key direction is the operationalization of the human firewall concept. Although this review identifies its core components—awareness, vigilance, behavioral compliance, and emotional regulation—there remains a need to translate these conceptual elements into measurable constructs within real-world environments.

Future work should focus on:

- Developing quantitative instruments to assess individual and organizational human firewall maturity.
- Conducting empirical validation of factors identified in this review across different demographic and technological contexts.
- Designing and evaluating intervention models that aim to enhance human resistance through training, feedback loops, and system-user integration.

This study provides a conceptual foundation for developing a detection system for social engineering attacks based on human characteristics. The subsequent research to be conducted by the author is designed to build the Social Engineering Attack Detection Model (SEADM), based on a human firewall approach within the context of mobile banking (m-banking) services. This model will integrate the psychological, social, and digital vulnerability factors identified in this study into a classification system and early detection of potential attacks.

IV Conclusion

This study was conducted to examine how individuals become victims of social engineering attacks and the factors that increase vulnerability to such attacks. Using the Systematic Literature Review (SLR) approach, 39 scholarly articles from various academic databases were thoroughly reviewed, covering the period from 2020 to 2024.

Based on the analysis of the collected literature, it was found that social engineering attacks tend to succeed not because of weaknesses in information technology systems, but due to vulnerabilities in the human aspect. Social engineering perpetrators exploit psychological responses and individual behavioral habits, such as excessive trust, panic, time pressure, and the habitual clicking of random links or digital information.

This study identifies five main patterns that describe how individuals can be manipulated into becoming victims, namely: emotional exploitation (such as fear or panic), abuse of authority, situational pressure, unawareness of digital risks, and cognitive fatigue due to information overload. These patterns do not occur randomly, but rather often follow a systematic and planned manipulative scheme.

Three main categories of individual vulnerability factors were also identified, namely:

- Psychological factors, such as blind trust, fear, or dependence on authority.
- Social factors, such as pressure from the work environment, a permissive organizational culture, and lack of security training.
- Digital-behavioral factors, including low digital literacy, password reuse, and the lack of conscious information verification.

These three categories interact and create an environment vulnerable to exploitation. Therefore, information security cannot be addressed solely from a technical perspective, but must be complemented with a human-centered approach.

Based on the findings of this study, the author provides the following recommendations:

1) *Human-Centered Security Approach:* Organizations should implement security strategies that focus on enhancing user awareness through training, simulations, and contextual education.

2) *Strengthening Digital Security Culture:* Information security values should be integrated into organizational policies and culture to create a work environment that is vigilant against social threats.

3) *Development of Human Firewall-Based Detection Models:* The findings of this study provide a foundation for further research to develop the Social Engineering Attack Detection Model (SEADM) in the context of mobile banking services.

4) *Enhancement of Public Digital Literacy:* Collective efforts are needed to raise digital security awareness among the public through educational campaigns, curricula, and cross-sector collaboration.

This study not only fills the gap in the literature by identifying patterns and factors of human vulnerability to social engineering, but also paves the way for the development of more adaptive and contextual security systems. The findings are expected to contribute to strengthening human-centered security systems and serve as a conceptual foundation for further applied and technical research.

REFERENCES

[1] D. K. Kumar, Shruthi, Sathwick, and Prathap, “Outcomes of social engineering: Understanding cybercriminals’ exploitation of human psychology,” *INTERANTIONAL J. Sci. Res. Eng. Manag.*, vol. 08, no. 11, pp. 1–7, Nov. 2024.

[2] S. W. Gabriel Bassett, C. David Hylender, Philippe Langlois, Alex Pinto, “Data Breach Investigations Report (DBIR),” 2022. doi: 10.1142/9789811218712_0009.

[3] Knowbe4, “Grid ® Report for Security Awareness Training | Summer 2023,” 2023.

[4] Z. Wang, H. Zhu, and L. Sun, “Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods,” *IEEE Access*, vol. 9, pp. 11895–11910, 2021, doi: 10.1109/ACCESS.2021.3051633.

[5] W. Syafitri, Z. Shukur, U. A. Mokhtar, R. Sulaiman, and M. A. Ibrahim, “Social Engineering Attacks Prevention: A Systematic Literature Review,” *IEEE Access*, vol. 10, pp. 39325–39343, 2022, doi: 10.1109/ACCESS.2022.3162594.

[6] S. Eftimie, R. Moinescu, and C. Racuciu, “Spear-Phishing Susceptibility Stemming From Personality Traits,” *IEEE Access*, vol. 10, pp. 73548–73561, 2022, doi: 10.1109/ACCESS.2022.3190009.

[7] K. F. Steinmetz, A. Pimentel, and W. R. Goe, “Performing social engineering: A qualitative study of information security deceptions,” *Comput. Human Behav.*, vol. 124, no. 106930, p. 106930, Nov. 2021, doi: 10.1016/j.chb.2021.106930.

[8] S. M. Albladi and G. R. S. Weir, “Predicting individuals’ vulnerability to social engineering in social networks,” *Cybersecurity*, vol. 3, no. 1, 2020, doi: 10.1186/s42400-020-00047-5.

[9] M. A. Siddiqi, W. Pak, and M. A. Siddiqi, “A Study on the Psychology of Social Engineering-Based Cyberattacks and Existing Countermeasures,” *Appl. Sci.*, vol. 12, no. 12, 2022, doi: 10.3390/app12126042.

[10] K. Chetioui, B. Bah, A. O. Alami, and A. Bahnasse, “Overview of Social Engineering Attacks on Social Networks,” *Procedia Comput. Sci.*, vol. 198, pp. 656–661, 2021, doi: 10.1016/j.procs.2021.12.302.

[11] N. Mashtalyar, U. N. Ntaganzwa, T. Santos, S. Hakak, and S. Ray, “Social Engineering Attacks: Recent Advances and Challenges,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12788 LNCS, Cham: Springer International Publishing, 2021, pp. 417–431. doi: 10.1007/978-3-030-77392-2_27.

[12] C. Sekhar Bhusal, “Systematic Review on Social Engineering: Hacking by Manipulating Humans,” *J. Inf. Secur.*, vol. 12, no. 01, pp. 104–114, 2021, doi: 10.4236/jis.2021.121005.

[13] H. Aldawood and G. Skinner, “Reviewing cyber security social engineering training and awareness programs-pitfalls and ongoing issues,” *Futur. Internet*, vol. 11, no. 3, 2019, doi: 10.3390/fi11030073.

[14] A. Alzahrani, “Coronavirus social engineering attacks: Issues and recommendations,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 5, pp. 154–161, 2020, doi: 10.14569/IJACSA.2020.0110523.

[15] H. Aldawood and G. Skinner, “An Advanced Taxonomy for Social Engineering Attacks,” *Int. J. Comput. Appl.*, vol. 177, no. 30, pp. 1–11, Jan. 2020, doi: 10.5120/ijca2020919744.

[16] Y. Choi, “Social Engineering Cyber Threats,” *J. Glob. Aware.*, vol. 4, no. 2, pp. 1–12, Dec. 2023, doi: 10.24073/jga/4/02/08.

[17] K. Saidi and Y. Prayudi, “Analisis Indikator Utama Dalam Information Security - Personality Threat Terhadap Phishing Attack,” *JUSTINDO (Jurnal Sist. dan Teknol. Inf. Indones.*, vol. 6, no. 1, pp. 21–30, Jun. 2021, doi: 10.32528/justindo.v6i1.3801.

[18] M. Jari, “A Comprehensive Survey of Phishing Attacks and Defences: Human Factors, Training and the Role of Emotions,” *Int. J. Netw. Secur. Its Appl.*, vol. 14, no. 5, pp. 11–24, Sep. 2022, doi: 10.5121/ijnsa.2022.14502.

[19] B. Coatesworth, “The psychology of social engineering,” *Cyber Secur. A Peer-Reviewed J.*, vol. 6, no. 3, p. 261, Mar. 2023, doi: 10.69554/aktg1392.

[20] A. Naz, M. Sarwar, M. Kaleem, M. A. Mushtaq, and S. Rashid, “A comprehensive survey on social engineering-based attacks on social networks,” *Int. J. Adv. Appl. Sci.*, vol. 11, no. 4, pp. 139–154, Apr. 2024, doi: 10.21833/ijaas.2024.04.016.

[21] F. A. Permana and A. Jamaludin, “Personal Data Vulnerability in the Digital Era: Study of Modus Operandi and Mechanisms to Prevent Phishing Crimes,” *J. Al-Hakim J. Ilm. Mahasiswa, Stud. Syariah, Huk. dan Filantr.*, vol. 5, no. 2, pp. 201–216, Nov. 2023, doi: 10.22515/jurnalalhakim.v5i2.7074.

[22] K. Sarpong Adu-Manu, R. Kwasi Ahiable, J. Kwame Appati, and E. Essel Mensah, “Phishing Attacks in Social Engineering: A Review,” *J. Cyber Secur.*, vol. 4, no. 4, pp. 239–267, 2022, doi: 10.32604/jcs.2023.041095.

[23] R. O. Oveh and G. O. Aziken, “Mitigating Social Engineering

Attack: A Focus on the Weak Human Link," in *Proceedings of the 5th International Conference on Information Technology for Education and Development: Changing the Narratives Through Building a Secure Society with Disruptive Technologies, ITED 2022*, Nov. 2022. doi: 10.1109/ITED56637.2022.10051202.

[24] T. Mokoena, T. Zuva, and M. Appiah, "Analysis of Social Engineering Attacks Using Exploit Kits," in *Advances in Intelligent Systems and Computing*, vol. 1224 AISC, Cham: Springer International Publishing, 2020, pp. 189–204. doi: 10.1007/978-3-030-51965-0_16.

[25] P. López-Aguilar and A. Solanas, "Human susceptibility to phishing attacks based on personality traits: The role of neuroticism," in *Proceedings - 2021 IEEE 45th Annual Computers, Software, and Applications Conference, COMPSAC 2021*, Jul. 2021, pp. 1363–1368. doi: 10.1109/COMPSAC51774.2021.00192.

[26] W. Fuertes *et al.*, "Impact of Social Engineering Attacks: A Literature Review," in *Developments and Advances in Defense and Security*, vol. 255, Singapore: Springer Singapore, 2022, pp. 25–35. doi: 10.1007/978-981-16-4884-7_3.

[27] I. Tjostheim and J. A. Waterworth, "Predicting personal susceptibility to phishing," in *Advances in Intelligent Systems and Computing*, vol. 1137 AISC, Cham: Springer International Publishing, 2020, pp. 564–575. doi: 10.1007/978-3-030-40690-5_54.

[28] M. R. Arabia-Obedoza, G. Rodriguez, A. Johnston, F. Salahdine, and N. Kaabouch, "Social Engineering Attacks A Reconnaissance Synthesis Analysis," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2020*, Oct. 2020, pp. 0843–0848. doi: 10.1109/UEMCON51285.2020.9298100.

[29] M. Mattera and M. M. Chowdhury, "Social Engineering: The Looming Threat," in *IEEE International Conference on Electro Information Technology*, May 2021, vol. 2021-May, pp. 56–61. doi: 10.1109/EIT51626.2021.9491884.

[30] P. Y. Leonov, A. V. Vorobyev, A. A. Ezhova, O. S. Kotelyanets, A. K. Zavalishina, and N. V. Morozov, "The main social engineering techniques aimed at hacking information systems," in *Proceedings - 2021 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBEREIT 2021*, May 2021, pp. 471–473. doi: 10.1109/USBEREIT51232.2021.9455031.

[31] A. S. V. Nair and R. Achary, "Social Engineering Defender (SE.Def): Human Emotion Factor Based Classification and Defense against Social Engineering Attacks," in *2023 International Conference on Artificial Intelligence and Applications, ICAIA 2023 and Alliance Technology Conference, ATCON-1 2023 - Proceeding*, Apr. 2023, pp. 1–5. doi: 10.1109/ICAIA57370.2023.10169678.

[32] S. A. Duman, R. Hayran, and I. Sogukpinar, "Impact Analysis and Performance Model of Social Engineering Techniques," in *ISDFS 2023 - 11th International Symposium on Digital Forensics and Security*, May 2023, pp. 1–6. doi: 10.1109/ISDFS58141.2023.10131771.

[33] A. Bishnoi, Garv, S. Bishnoi, and N. Gupta, "Comprehensive Assessment of Reverse Social Engineering to Understand Social Engineering Attacks," in *Proceedings - 5th International Conference on Smart Systems and Inventive Technology, ICSSIT 2023*, Jan. 2023, pp. 681–685. doi: 10.1109/ICSSIT55814.2023.10061054.

[34] S. Longtchi, Theodore Tangie and Rodriguez, Rosana Montañez and Al-Shawaf, Laith and Atyabi, Adham and Xu, "Internet-Based Social Engineering Psychology, Attacks, and Defenses: A Survey," *Proc. IEEE*, vol. 112, no. 3, pp. 210–246, 2024, doi: 10.1109/JPROC.2024.3379855.

[35] T. V. Tulupieva, "Psychological Aspects of the Organization's Information Security in the Context of Socio-engineering Attacks," *Adm. Consult.*, no. 2, pp. 123–128, Mar. 2022, doi: 10.22394/1726-1139-2022-2-123-138.

[36] R. Ribeiro, N. Mateus-Coelho, and H. Mamede, "Improving Social Engineering Resilience In Enterprises," *ARIS2 - Adv. Res. Inf. Syst. Secur.*, vol. 3, no. 1, pp. 34–65, Aug. 2023, doi: 10.56394/aris2.v3i1.30.

[37] N. AKYEŞİLMEN and A. ALHOSBAN, "Non-Technical Cyber-Attacks and International Cybersecurity: The Case of Social Engineering," *Gaziantep Univ. J. Soc. Sci.*, vol. 23, no. 1, pp. 342–360, Jan. 2024, doi: 10.21547/jss.1346291.

[38] K. Mahanta and H. B. Maringanti, "Social engineering attacks and countermeasures," in *Perspectives on Ethical Hacking and Penetration Testing*, IGI Global, 2023, pp. 307–337. doi: 10.4018/978-1-6684-8218-6.ch013.

[39] K. Thakur, J. Shan, and A. S. K. Pathan, "Innovations of phishing defense: The mechanism, measurement and defense strategies," *Int. J. Commun. Networks Inf. Secur.*, vol. 10, no. 1, pp. 19–27, Apr. 2018, doi: 10.17762/ijcnis.v10i1.2991.

[40] J. Vargis and D. Murphy, "Mitigating Social Engineering Attacks on the Elderly: Personalized Countermeasures to Enhance Cyber Situational Awareness," in *The European Conference on Aging & Gerontology 2023: Official Conference Proceedings*, Sep. 2023, pp. 43–53. doi: 10.22492/issn.2435-4937.2023.5.

[41] A. Şandor, G. Tont, and E. Simion, "A Mathematical Model for Risk Assessment of Social Engineering Attacks," *TEM J.*, vol. 11, no. 1, pp. 334–338, Feb. 2022, doi: 10.18421/TEM111-42.

[42] K. F. Steinmetz, "The Identification of a Model Victim for Social Engineering: A Qualitative Analysis," *Vict. Offenders*, vol. 16, no. 4, pp. 540–564, May 2021, doi: 10.1080/15564886.2020.1818658.