

Cognitive Biases as Predictors of Digital Fraud Victimization Among Retail Banking Customers

D M Renuka Devi¹

Abstract

Digital banking has expanded access to financial services while also creating communication environments in which fraudulent requests can exploit rapid judgment, urgency, familiarity, and perceived loss.

This study examined whether confirmation bias, anchoring bias, loss aversion, and availability heuristic were associated with self-reported digital-fraud victimization among retail banking customers.

Cross-sectional predictive-correlational design was used with 123 complete survey records. Cognitive-bias composites were calculated as item means, and victimization was operationalized as the sum of 11 frequency-coded attempted or successful fraud incidents reported for the preceding 12 months. Descriptive statistics, internal-consistency reliability, Pearson correlations, hierarchical multiple regression with HC3 robust standard errors, negative-binomial sensitivity analysis, demographic comparisons, and exploratory moderation analyses were conducted.

Cognitive-bias composites demonstrated good to excellent internal consistency ($\alpha = .838-.921$). Victimization was most strongly correlated with loss aversion ($r = .461$) and anchoring ($r = .385$). Adding the four biases to demographic controls increased explained variance from $R^2 = .127$ to $R^2 = .327$ ($\Delta R^2 = .200$). In the adjusted model, anchoring ($B = 2.505$, $p = .037$) and loss aversion ($B = 2.978$, $p = .002$) were positive predictors, whereas confirmation bias and availability heuristic were not statistically significant unique predictors. Negative-binomial sensitivity analysis showed convergent positive associations for anchoring (IRR = 1.345) and loss aversion (IRR = 1.503).

Submitted : October 10, 2026

Published : October 31, 2026

DOI: doi.org/10.65320/jce.vol.1.issue5.53

1* Psychology Scholar, Jain University, Bengaluru, India.
(renuka097@gmail.com)

**Corresponding Author*



This paper is blockchain-verified. Scan the QR code to verify its authenticity.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



© 2026 The Author(s). Journal of Cortexplore published by Managers Without Borders Community Confederation.

Support a qualified, non-causal association between selected cognitive tendencies and reported fraud burden and indicate that fraud-prevention design may benefit from interventions that interrupt salient anchors and loss-framed urgency.

Keywords: Cognitive Bias | Digital Fraud | Phishing | Anchoring | Loss Aversion | Retail Banking

I. INTRODUCTION

1.1 Digital finance as a human decision environment

The contemporary retail-banking environment is increasingly mediated by digital interfaces. Customers can move between mobile banking applications, internet banking, UPI services, QR-code payments, SMS, email, messaging applications, and telephone channels within a single financial episode. These systems reduce transaction friction, but the same speed and convenience can reduce the time available for deliberate verification. The Reserve Bank of India has continued to identify phishing, vishing, malware, skimming, and related cyber risks as important components of the digital-payment environment and has emphasized customer protection alongside technological development (Reserve Bank of India, 2025). This context makes digital fraud a problem of both information security and human judgment.

A fraudulent communication rarely depends only on a technical vulnerability. Many schemes are structured around an interaction in which the recipient must interpret a message, judge its legitimacy, decide whether a requested action is safe, and often respond under time pressure. The attacker may imitate a trusted institution, present a familiar transaction, display a numerical amount, claim that an account will be frozen, or suggest that immediate verification is required. The customer is therefore not simply confronted with a cybersecurity warning; the customer is placed in a decision environment deliberately constructed to influence attention and action.

Information-processing research provides a useful basis for understanding this environment. Models of phishing susceptibility have shown that recipients may rely on relatively simple cues rather than systematically examining every available feature of a communication. Vishwanath et al. (2011) linked individual differences in phishing vulnerability to an integrated information-processing model, and later work emphasized suspicion, cognition, and automaticity in phishing susceptibility (Vishwanath et al., 2018). Luo et al. (2013) used the Heuristic–Systematic Model (HSM) to explain how cognitive processing may contribute to phishing victimization. More recent synthesis continues to indicate that susceptibility is heterogeneous and that the evidence for particular predictors varies across contexts and outcome definitions (Spicer & Nurse, 2026).

The present study builds on this literature by examining four cognitive tendencies together rather than treating them as isolated explanations. Confirmation bias concerns the tendency to give greater weight to information that fits an existing expectation (Nickerson, 1998). Anchoring concerns disproportionate reliance on an initial value or cue. Loss aversion concerns the comparatively strong psychological weight of potential losses. The availability heuristic concerns judgments influenced by how readily relevant examples or information come to mind (Kahneman & Frederick, 2002; Tversky & Kahneman, 1973). These mechanisms are conceptually distinct, but in a fraud encounter they can be activated by the same message.

1.2 Cognitive mechanisms relevant to fraud

Confirmation bias may become relevant when a customer forms an early judgment that a communication is genuine. A familiar bank name, recognizable visual identity, plausible transaction description, or expected communication channel can provide confirming information. Once legitimacy has been accepted provisionally, contradictory details may receive less attention. This mechanism does not imply that a person deliberately ignores evidence; rather, the interpretation of subsequent evidence may be influenced by the initial expectation (Nickerson, 1998).

Anchoring is especially relevant to digital financial interactions because these interactions contain abundant numerical and sequential information. Account balances, transaction amounts, transfer limits, beneficiary numbers, refund values, payment deadlines, and countdown messages can all function as salient reference points. A fraudulent message may place an initial number or label at the beginning of the decision sequence and encourage subsequent judgments relative to that cue. Research on contextualized phishing has similarly emphasized that message context and presentation can affect how recipients respond to deceptive communications (Hassandoust et al., 2019).

Loss aversion provides a different pathway. Fraud messages frequently emphasize what the customer may lose if action is delayed: access to an account, a payment, a refund, a card, a service, or money. Under such framing, verification can be made to feel like a threat-management task rather than a routine security step. Dual-process accounts of fraud argue that urgency and emotional pressure can increase reliance on rapid processing and reduce the opportunity for corrective analysis (Xu et al., 2025). Experimental work has also shown that time pressure and deception cues influence responses to phishing communications (Butavicius et al., 2022).

Availability can influence the interpretation of a new message through the accessibility of prior

examples. A recently encountered fraud story, a vivid news report, a memorable transaction, or a familiar scam narrative may make particular interpretations easier to retrieve. However, availability is directionally ambiguous. A vivid fraud example can increase caution, but it can also make a deceptive narrative feel familiar. Theoretical plausibility therefore does not guarantee an independent empirical association with victimization.

1.3 Why victimization frequency requires careful definition

An important issue in the fraud literature is the distinction between susceptibility, compliance, exposure, victimization, and financial loss. Experimental phishing studies may measure whether a person clicks a simulated link or identifies a deceptive message. Other studies measure intention to comply, detection accuracy, or self-reported experience. Administrative data may capture reported financial losses. These outcomes are related but not interchangeable. A person may encounter many fraudulent messages and successfully resist each one, whereas another person may encounter fewer messages but experience a substantial financial loss.

The present study therefore defines its outcome narrowly. Digital-fraud victimization is represented by the summed frequency of 11 specified attempted or successful incident types reported over the preceding 12 months. Each incident is coded from 0 to 3, yielding a theoretical range of 0–33. The separate monetary-loss item is not included in this index because rupee loss and incident frequency represent different quantities. This distinction is important for both interpretation and reproducibility.

1.4 Research gap and study objectives

Three gaps motivated the study. First, a substantial portion of phishing and social-engineering research focuses on broad constructs such as awareness, self-efficacy, personality, or susceptibility, whereas fewer studies

simultaneously estimate several domain-specific cognitive mechanisms. Second, experimental susceptibility measures do not necessarily represent accumulated real-world incident burden. Third, there remains value in examining these mechanisms within an Indian retail-banking environment characterized by extensive use of UPI, mobile banking, messaging, QR payments, and telephone-based financial communication.

The study consequently asked five questions: what are the levels and interrelationships of the four cognitive-bias constructs; which fraud incidents are most frequently reported; do the four biases jointly and uniquely predict victimization frequency; do demographic characteristics relate to cognitive-bias profiles and victimization burden; and do age, gender, banking experience, and digital literacy alter bias–victimization associations? The primary hypothesis was that the four cognitive biases would jointly predict victimization. Additional analyses examined demographic differences and exploratory moderation.

The study was designed as an empirical contribution rather than as a diagnostic assessment of individuals. A cognitive bias is treated as a decision tendency, not as a stable defect or a sufficient explanation for fraud. This framing is important because fraud is produced by an interaction among attacker behavior, message design, technological affordances, exposure opportunities, and human judgment. The objective is therefore to identify potentially useful psychological mechanisms without transferring responsibility for fraud entirely to customers.

II. METHODOLOGY

2.1 Design and participants

The study used a quantitative, cross-sectional, predictive-correlational design. The target population was adult retail banking customers who used digital banking platforms and had at least six months of digital-banking experience. Recruitment was purposive and voluntary. The planned sample size was 200; 123 complete records were available for analysis. All 123 records

contained complete responses for the cognitive-bias measures and incident-frequency items, and no duplicate complete rows were identified.

The achieved sample should therefore be understood as a complete analytic sample below the original recruitment target rather than as a statistically representative sample of retail banking customers in India. The distinction matters because data completeness and sampling representativeness are different properties. The absence of item-level missingness simplifies scoring, but purposive or voluntary recruitment limits population-level generalizability.

2.2 Measures and operationalization

The questionnaire comprised demographic and banking information, four cognitive-bias sections, and a digital-fraud section. Confirmation bias was measured with eight Likert-type items, anchoring bias with six items, loss aversion with seven items, and availability heuristic with five items. Responses used a five-point agreement scale. Composite scores were calculated as arithmetic item means so that all four constructs remained on the original 1–5 response metric.

The victimization measure comprised 11 frequency-coded incident items. Participants reported how often each specified incident had occurred during the preceding 12 months using a 0–3 coding scheme: none, once, two to three times, or more than three times. The 11 item scores were summed, producing an index from 0 to 33. The questionnaire also contained a separate monetary-loss question. That item was deliberately excluded from the frequency index because adding a rupee amount to incident counts would produce an uninterpretable composite.

An important documentation issue was retained transparently. Earlier methodology wording referred to a 12-item victimization measure, whereas inspection of the administered questionnaire showed 11 frequency-coded incident items followed by a separate financial-loss item. The analysis followed the observable

questionnaire structure rather than forcing the monetary-loss question into the frequency measure. This distinction should be retained in future versions of the instrument.

No reverse-scored cognitive-bias items were identified in the administered instrument, so reverse coding was not introduced retrospectively. Cronbach's alpha was used to assess internal consistency. The resulting coefficients are reported as reliability indicators and are not treated as evidence of factorial validity. The available data were not sufficient for a full psychometric validation of the instrument.

2.3 Procedure and ethics

Participants completed an anonymous survey after consent. The source methodology states that ethical approval was secured, participation was voluntary, informed consent was obtained, anonymity and data-security procedures were used, and participants could withdraw. The manuscript retains these statements subject to verification against the institutional ethics documentation before journal submission. No personally identifying participant information is reported in the manuscript.

2.4 Statistical analysis

Descriptive statistics were used to characterize the sample, cognitive-bias composites, and victimization index. Pearson correlations were used to examine associations among the continuous study variables. Internal consistency was evaluated with Cronbach's alpha.

Hierarchical multiple regression was used for the primary predictive question. Demographic controls were entered first, followed by the four cognitive-bias composites. The change in R^2 was used to evaluate the incremental contribution of the cognitive-bias block. Because the victimization outcome was highly dispersed and ordinary least-squares assumptions were imperfect, HC3 heteroskedasticity-consistent standard errors were used for coefficient inference. This approach

retains the prespecified regression framework while reducing sensitivity to heteroskedasticity.

A negative-binomial regression was estimated as a sensitivity analysis because the victimization index was count-like and substantially overdispersed. The variance-to-mean ratio was 12.11. The negative-binomial model was not used to replace the primary analysis or to select a more favorable result; rather, it was used to examine whether the principal associations remained directionally similar under a model more appropriate for overdispersed counts.

Demographic comparisons were examined using analysis of variance or equivalent group comparisons, with effect sizes considered alongside p values. Exploratory interactions between each cognitive bias and age, gender, banking experience, and digital literacy were also examined. Sixteen interactions were tested. Because the sample size was modest and interaction estimates are typically imprecise, the moderation analysis was treated as exploratory rather than confirmatory.

Statistical significance was evaluated at $\alpha = .05$, but interpretation emphasized effect magnitude, confidence intervals, model structure, and the distinction between association and causation. In particular, a statistically significant regression coefficient was not interpreted as evidence that a cognitive bias caused a fraud incident

III. RESULTS

3.1 Participant characteristics

The final sample comprised 123 retail banking customers. The largest age group was 25–34 years ($n = 49$, 39.8%), followed by 18–24 years ($n = 32$, 26.0%). Participants aged 45–54 years accounted for 14.6%, those aged 35–44 years for 10.6%, and those aged 55 years or older for 8.9%. The gender distribution was relatively balanced: 63 participants were female (51.2%), 59 were male (48.0%), and one preferred not to say (0.8%). Most participants reported either a bachelor's degree (41.5%) or master's degree (46.3%). Monthly income was distributed across four categories,

with 34.1% reporting less than ₹50,000 and 17.1% reporting more than ₹200,000. Mean banking experience was 59.0 months (SD = 37.18), with a median of 48 months. Mean self-rated digital literacy was 4.07 on a five-point scale (SD = 0.91).

Table 1

Characteristic	Distribution
Age	18–24: 32 (26.0%); 25–34: 49 (39.8%); 35–44: 13 (10.6%); 45–54: 18 (14.6%); 55+: 11 (8.9%)
Gender	Female: 63 (51.2%); Male: 59 (48.0%); Prefer not to say: 1 (0.8%)
Education	High School: 5 (4.1%); Bachelor's: 51 (41.5%); Master's: 57 (46.3%); PhD/Professional: 10 (8.1%)
Income	<₹50,000: 42 (34.1%); ₹50,001–100,000: 38 (30.9%); ₹100,001–200,000: 22 (17.9%); >₹200,000: 21 (17.1%)
Banking experience	M = 59.0 months; SD = 37.18; median = 48; range = 6–180
Digital literacy	M = 4.07/5; SD = 0.91; median = 4

Participant characteristics

Note. N = 123. Percentages may not sum to exactly 100 because of rounding.

3.2 Descriptive statistics and reliability

Availability heuristic had the highest mean among the four cognitive-bias composites (M = 3.45, SD = 1.09). Confirmation bias had a mean of 2.64 (SD = 0.93), while anchoring bias and loss aversion both had means of 2.57, with SDs of 1.05 and 1.19, respectively. All four composites showed good to excellent internal consistency: confirmation bias α = .838, anchoring bias α = .872, loss aversion α = .921, and availability heuristic α = .873.

The victimization index had a mean of 7.08 (SD = 9.26), a median of 3, and a range from 0 to 33. The

difference between the mean and median, together with the variance-to-mean ratio of 12.11, indicates substantial dispersion and a pronounced upper tail. The outcome should therefore not be interpreted as approximately symmetric or as a conventional continuous psychological scale.

Table 2

Variable	M	SD	Median	α	Victimization r	Range
Confirmation bias	2.64	.93	2.38	.838	.178	1–5
Anchoring bias	2.57	1.05	2.67	.872	.385	1–5
Loss aversion	2.57	1.19	2.43	.921	.461	1–5
Availability heuristic	3.45	1.09	3.60	.873	.268	1–5
Victimization index	7.08	9.26	3.00	—	1.000	0–33

Descriptive statistics, reliability, and zero-order association with victimization

Note. Cronbach's α is reported for the four cognitive-bias composites. Victimization is the sum of 11 frequency-coded incident items. Pearson r values show the zero-order association with the victimization index.

3.3 Digital-fraud incident profile

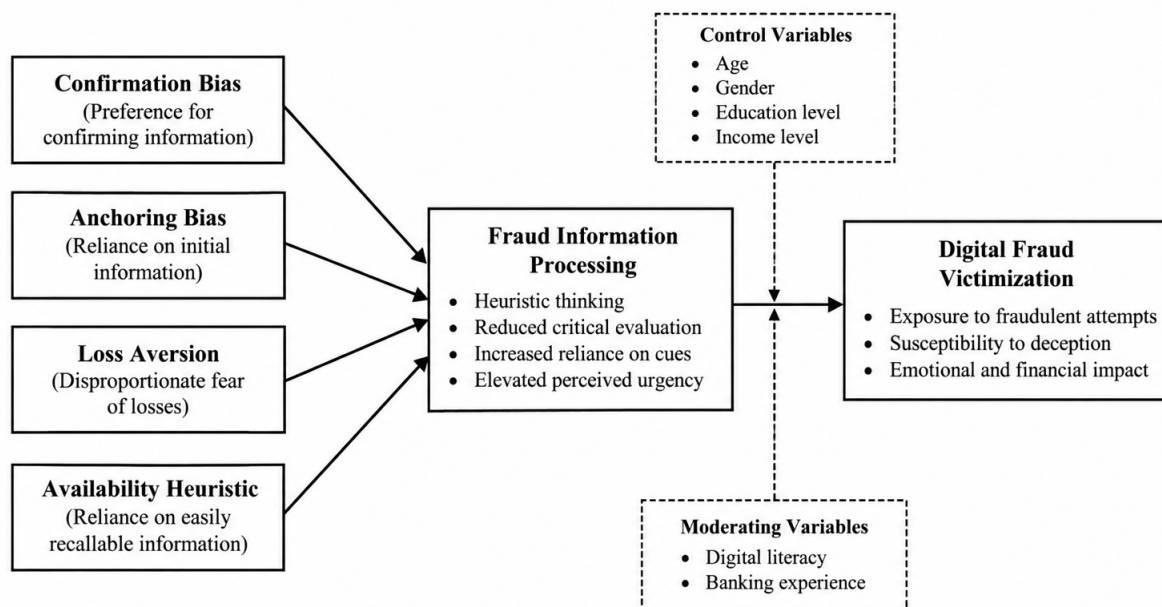
The most frequently reported incident category was phishing email or SMS leading to link clicking or information sharing, reported by 62 participants (50.4%). Social-engineering calls pretending to be bank officials were reported by 58 participants (47.2%), followed by investment scams using WhatsApp or banking pretexts (39.0%). Email attachments presented as bank support communications were reported by 37.4%, refund fraud requiring a verification payment by 36.6%, and false beneficiaries or merchants leading to money transfer by 31.7%.

Other reported categories included account compromise through public Wi-Fi or online shopping, fake bank applications or malware, ATM skimming or phishing, unauthorized UPI transactions after OTP disclosure, and QR-code scams. Across the 11 listed incident types, 90 of the 123 participants (73.2%) reported at least one listed attempted or successful incident. This figure must not be described as a successful-fraud prevalence estimate because the questionnaire combines attempted and successful experiences.

The incident profile is important for interpreting the psychological findings. Several of the most

common categories are communication-intensive and therefore compatible with mechanisms involving urgency, initial cues, authority, familiarity, or threat. The data do not establish which cognitive mechanism operated in any particular incident, but the distribution shows that the outcome reflects a heterogeneous set of digital-fraud encounters rather than a single attack channel.

Figure 1. *Conceptual Framework of Cognitive Biases and Digital-Fraud Victimization*



Note. The framework summarizes the study variables and the hypothesized analytical relationship between the four cognitive-bias constructs and the victimization outcome. Arrows represent analytical relationships, not causal effects.

3.4 Correlations among cognitive biases and victimization

The four cognitive-bias composites were positively correlated with one another. The strongest association was between confirmation bias and anchoring bias ($r = .712$), followed by anchoring and loss aversion ($r = .628$), loss aversion and availability ($r = .568$), confirmation and loss

aversion ($r = .500$), anchoring and availability ($r = .452$), and confirmation and availability ($r = .375$).

Victimization correlated most strongly with loss aversion ($r = .461$), followed by anchoring bias ($r = .385$), availability heuristic ($r = .268$), and confirmation bias ($r = .178$). These zero-order relationships indicate that the four constructs were not interchangeable and that the strongest bivariate relationships with victimization involved loss-related and anchoring-related tendencies. At the same time, the high correlation between confirmation and anchoring indicates meaningful shared variance, making isolated interpretation of multivariable coefficients inappropriate without considering suppression or overlapping explanatory content.

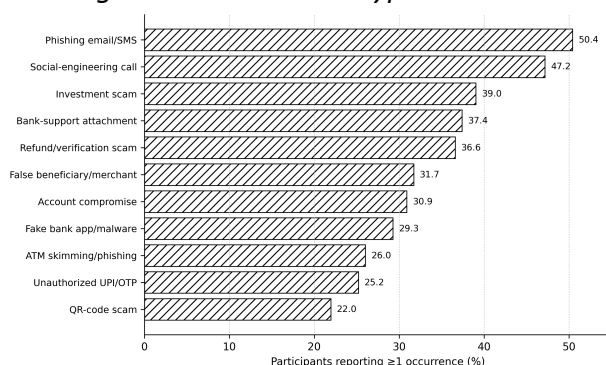
3.5 Hierarchical regression

The demographic-control model produced $R^2 = .127$ and adjusted $R^2 = .014$, with $F = 1.12$ and $p = .349$. After adding the four cognitive-bias composites, R^2 increased to $.327$ and adjusted R^2 to $.210$. The full model was statistically significant ($F = 2.08$, $p < .001$), representing an incremental R^2 of $.200$ attributable to the bias block within the specified model.

The adjusted coefficient estimates showed two statistically significant positive predictors. Anchoring bias had $B = 2.505$, HC3 SE = 1.198, $t = 2.09$, $p = .037$, with a 95% CI of [0.156, 4.853]. Loss aversion had $B = 2.978$, HC3 SE = 0.940, $t = 3.17$, $p = .002$, with a 95% CI of [1.137, 4.820]. Availability heuristic was not statistically significant ($B = -0.068$, $p = .913$), and confirmation bias was also not statistically significant in the adjusted model ($B = -2.059$, $p = .105$).

The negative sign of the adjusted confirmation coefficient should not be interpreted as a protective effect. Its zero-order correlation with victimization was positive ($r = .178$), while its correlation with anchoring was strong ($r = .712$). The reversal is therefore more appropriately interpreted as a shared-variance or suppression phenomenon requiring replication. The coefficient answers the conditional question of association after the other predictors have been included, not the broader question of whether confirmation-consistent processing is generally beneficial or harmful.

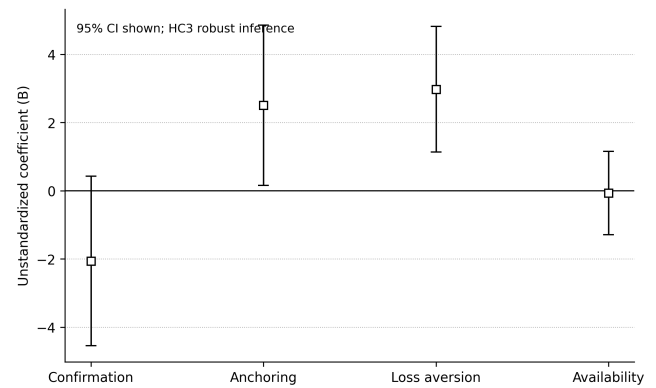
Figure 2. Percentage of Participants Reporting Each Digital-Fraud Incident Type



Note. Percentages represent participants reporting at least one listed attempted or

successful incident during the preceding 12 months. The measure does not estimate successful-fraud prevalence.

Figure 3. Adjusted Regression Coefficients for the Four Cognitive Biases



Note. Points represent unstandardized regression coefficients and error bars represent 95% confidence intervals. HC3 robust standard errors were used for inference.

3.6 Negative-binomial sensitivity analysis and diagnostics

The negative-binomial sensitivity analysis showed convergence for anchoring and loss aversion. Anchoring had an incidence-rate ratio of 1.345 ($p = .027$, 95% CI [1.035, 1.748]), indicating a positive association with expected victimization frequency. Loss aversion had IRR = 1.503 ($p < .001$, 95% CI [1.198, 1.887]). Confirmation bias had IRR = .687 ($p = .035$, 95% CI [.485, .975]), while availability heuristic had IRR = 1.019 ($p = .871$, 95% CI [.808, 1.286]). The count-model results therefore broadly reinforced the positive associations of anchoring and loss aversion while confirming uncertainty around availability. The confirmation estimate again requires cautious interpretation because its direction differs from the bivariate relationship and from the conceptual expectation.

The diagnostic audit showed a variance-to-mean ratio of 12.11, supporting the decision to conduct a count-model sensitivity analysis. The Breusch–Pagan test was not statistically significant ($p = .128$), whereas Jarque–Bera and Shapiro–Wilk tests indicated non-normality (both $p < .001$). The maximum variance inflation factor was 2.85, indicating that the four cognitive-bias predictors were correlated but remained below the prespecified VIF threshold of 3.0.

Table 3. Primary regression and negative-binomial sensitivity estimates

Predictor	B / IRR	HC 3 SE	p	95% CI	Interpretation
Confirmation bias	B = -2.059	1.269	.105	[-4.545, 0.428]	OLS unique association not significant
Anchoring bias	B = 2.505	1.198	.037	[0.156, 4.853]	Positive adjusted association
Loss aversion	B = 2.978	.940	.002	[1.137, 4.820]	Positive adjusted association
Availability heuristic	B = -0.068	.622	.913	[-1.287, 1.151]	OLS unique association not significant
Anchoring sensitivity	IRR = 1.345	—	.027	[1.035, 1.748]	Positive count-model association
Loss-aversion sensitivity	IRR = 1.503	—	<.001	[1.198, 1.887]	Positive count-model association

Note. The primary model controlled for demographic variables and included all four cognitive-bias composites. HC3 robust standard errors were used. IRR = incidence-rate ratio. The negative-binomial estimates are sensitivity analyses, not replacements for the primary model.

3.7 Demographic differences and exploratory moderation

The four cognitive-bias composites did not show statistically significant categorical differences across age, gender, education, or income. Banking experience had modest negative

correlations with loss aversion ($r = -.219$, $p = .015$) and availability heuristic ($r = -.185$, $p = .041$). These associations are small and should not be interpreted as evidence that experience eliminates cognitive bias.

Victimization displayed a different pattern. Continuous victimization burden differed by gender, with males reporting a higher mean index than females (approximately 9.83 versus 4.62; $p = .002$ in the continuous comparison). The corresponding omnibus gender test was $F = 5.50$, $p = .005$, $\eta^2 = .084$. Age, education, and income did not show statistically significant differences in the continuous victimization outcome. Binary comparisons of whether a participant reported at least one incident were not statistically significant across the demographic categories. Thus, the gender result appears more closely related to the amount or frequency of reported incident burden than to a simple distinction between any incident versus none.

Sixteen bias-by-moderator interactions were examined across age, gender, banking experience, and digital literacy. None crossed $p < .05$; the smallest p value was approximately .053. Given the sample size and the number of interaction terms, these results are best characterized as inconclusive rather than as evidence that moderation is absent.

IV. DISCUSSION

4.1 Principal findings

This study examined whether four cognitive-bias constructs were associated with self-reported digital-fraud victimization among retail banking customers. The central result is that the cognitive-bias block contributed meaningful additional explanatory variance beyond demographic controls, increasing R^2 from .127 to .327. Within the adjusted model, anchoring and loss aversion showed positive associations with victimization frequency, and those associations remained positive in the negative-binomial sensitivity analysis. Confirmation bias and availability heuristic did not demonstrate statistically

significant unique positive effects in the principal model.

The results support a qualified interpretation of the primary hypothesis. The four cognitive mechanisms considered jointly were associated with variation in the victimization index, but the evidence does not support the claim that all four are independent predictors of fraud burden. In particular, the data point toward anchoring and loss aversion as the mechanisms that retained the clearest adjusted associations in this sample. This distinction is more informative than treating the four constructs as a single undifferentiated measure of cognitive vulnerability.

4.2 Anchoring and the role of initial cues

Anchoring was positively related to victimization both before and after adjustment. The bivariate correlation was $r = .385$, and the adjusted regression coefficient was $B = 2.505$. The negative-binomial IRR of 1.345 provided a convergent count-model estimate. Although these findings are not causal, they are compatible with a process in which an initial value or cue influences later evaluation of a digital-finance communication.

The relevance of anchoring is plausible in modern banking interfaces because customers encounter numerical information continuously. A scammer can present an amount, beneficiary identifier, transaction reference, refund value, deadline, or countdown before the customer has independently established whether the communication is legitimate. Once an initial value becomes psychologically salient, later information may be evaluated relative to that reference point. Contextualized phishing research similarly indicates that the presentation and situational framing of deceptive messages can influence responses (Hassandoust et al., 2019).

The present data cannot determine which particular anchors were most influential. The questionnaire includes several realistic situations, but it does not record the cognitive sequence of an actual fraud episode. A high anchoring score could reflect a general tendency, while an observed victimization event could have been driven by a

different mechanism. Future research should therefore move from self-report tendency measures toward behavioral experiments that manipulate the order, magnitude, and salience of financial cues.

4.3 Loss aversion and threat-framed urgency

Loss aversion showed the strongest bivariate relationship with victimization and the largest positive adjusted coefficient. The negative-binomial estimate was also positive and statistically significant. This convergence across model families is notable because the models make different distributional assumptions.

A plausible interpretation is that fraud messages can convert routine security decisions into perceived emergencies. Instead of asking the customer to decide whether a message is trustworthy, the communication may imply that inaction will produce an immediate loss: account suspension, card blocking, penalty, refund cancellation, temporary hold, or loss of access. The customer may consequently perceive verification as a costly delay. The attacker benefits when the perceived cost of waiting exceeds the perceived cost of complying.

This interpretation fits dual-process accounts in which urgency and emotional pressure can increase reliance on rapid processing (Xu et al., 2025). It also aligns with experimental evidence that time pressure and deception cues can influence phishing responses (Butavicius et al., 2022). However, the current study cannot establish that loss-framed messages caused the reported incidents. The questionnaire measures a tendency and an accumulated outcome, not the psychological state present at the moment of an actual attack.

The practical implication is therefore not simply to tell customers to be less loss averse. A more useful prevention approach is to alter the decision environment. Bank interfaces can explicitly state that legitimate institutions do not require immediate action through unsolicited links or calls, introduce deliberate confirmation steps for high-risk transactions, and phrase warnings so that

the safer behavior does not feel like a threat to access or money. The objective is to reduce the effectiveness of threat-based urgency rather than to place responsibility for resisting it solely on the customer.

4.4 Confirmation bias and shared variance

Confirmation bias showed a positive zero-order relationship with victimization but a negative, non-significant adjusted coefficient. This pattern is analytically important. The strong correlation between confirmation bias and anchoring suggests that the constructs share substantial variance. When both are included in the same regression equation, the coefficient for confirmation represents the association remaining after the shared component with anchoring and the other predictors has been statistically removed.

It would therefore be inappropriate to conclude that confirmation bias protects customers from fraud. The direction reversal is more consistent with suppression or shared-variance effects, especially because the confidence interval includes zero. This result illustrates why multivariable evidence should not be reduced to a simple list of significant and non-significant *p* values. A theoretically plausible construct can show a positive bivariate association but fail to retain an independent coefficient when correlated mechanisms are modeled simultaneously.

The finding also has implications for measurement. The reliability coefficient for confirmation bias was good, but reliability does not demonstrate that the questionnaire items isolate a single latent construct. Some confirmation items may naturally overlap with trust, familiarity, or anchoring-related decision tendencies. Future psychometric work should test factor structure and discriminant validity using larger samples before strong claims are made about the separability of these constructs.

4.5 Availability heuristic and contextual directionality

Availability showed a positive bivariate correlation with victimization but no independent effect in

either the adjusted OLS model or the negative-binomial sensitivity analysis. This does not demonstrate that availability is irrelevant to fraud. Instead, it suggests that general accessibility of examples did not explain additional victimization variation after the other cognitive measures and controls were considered.

The null adjusted result may reflect several possibilities. First, the availability items may capture a broad memory-based tendency that is less proximal to fraud behavior than the message-level cues represented by anchoring and loss framing. Second, availability can operate in opposite directions. A vivid fraud story can make a person more cautious, but it can also make a new scam narrative feel familiar. Third, the construct may interact with the specific content of the available information. The present survey did not measure whether the examples accessible to a respondent were cautionary, reassuring, or misleading.

Future research could address this issue experimentally by manipulating the salience of different types of prior information before participants evaluate standardized banking communications. Such designs would allow researchers to distinguish availability that increases perceived threat from availability that increases perceived familiarity.

4.6 Incident profile and the human attack surface

The incident profile shows that the sample encountered a range of communication-based and transaction-based threats. Qualitative research on investment-scam victims similarly illustrates how financial decision-making can be affected by anchoring, emotion, and social influence in scam contexts (Cuadra et al., 2025). Phishing email or SMS, social-engineering calls, investment scams, fraudulent attachments, and refund fraud were among the more commonly reported categories. The prevalence of these channels is consistent with the broader understanding that fraud can move across technological and social boundaries. A customer may begin with an SMS, continue through a web page, receive a telephone call, and

complete a transaction through a banking or UPI application.

This channel switching matters for psychological theory because cognitive processing does not occur independently of interface design. The same customer may respond differently when a request arrives through an expected banking application compared with an unexpected phone call. Contextual specificity can increase the plausibility of a deceptive communication, and research has shown that self-predictions of how people will respond do not always correspond to observed behavior (Hassandoust et al., 2019). The present incident measure cannot isolate these contextual differences, but it reinforces the need to examine fraud as an interaction rather than as a single static stimulus.

The finding that 73.2% of participants reported at least one listed attempted or successful incident should be interpreted carefully. It is a sample-specific indicator of exposure to the listed categories, not an estimate of successful fraud prevalence in the banking population. Because the questionnaire combines attempts and successful events, a participant may have encountered a fraudulent message and resisted it. This distinction is essential when translating survey results into policy recommendations.

4.7 Demographic findings and the limits of demographic profiling

The absence of significant categorical differences in the four cognitive-bias composites suggests that the measured tendencies were not strongly stratified by age, gender, education, or income in this sample. Banking experience showed modest negative associations with loss aversion and availability, but the effects were not large enough to support a simple claim that experience protects against cognitive vulnerability.

Victimization burden differed by gender in the continuous outcome, with males reporting higher mean scores. Yet binary any-incident comparisons were not significant. This distinction suggests that demographic variables may relate differently to incident frequency and incident occurrence.

Exposure opportunities, transaction volume, occupational activity, reporting behavior, and unmeasured contextual factors could contribute to the observed difference. The data do not justify interpreting gender as a causal explanation of fraud vulnerability.

This pattern is consistent with recent phishing research indicating that demographic and technology-competence variables do not always predict susceptibility consistently, while more proximal psychological factors such as detection self-efficacy can matter (Ribeiro et al., 2024). The broader implication is that fraud prevention should avoid demographic stereotypes. A customer should not be assumed to be safe or unsafe simply because of age, education, income, or gender. Situational safeguards can be applied more consistently and with less risk of excluding or stigmatizing groups.

4.8 Theoretical contribution

The findings support an integrated view in which cognitive biases are not independent psychological switches but overlapping tendencies that may become consequential under particular communication conditions. The HSM provides a useful framework for this interpretation because it emphasizes the relationship between heuristic processing, systematic evaluation, motivation, prior knowledge, and message cues (Luo et al., 2013). The current results suggest that some cue-sensitive mechanisms may be more closely associated with accumulated fraud burden than others.

A conceptual sequence can be proposed cautiously. A fraudulent message may first establish an initial cue or reference point. Confirmation-consistent interpretation may then reduce attention to contradictory information. A threatened loss may increase urgency and make delay feel costly. Accessible memories or familiar narratives may influence plausibility. The resulting response is also shaped by the transaction context, channel, prior experience, and available time. The present cross-sectional data cannot test this sequence directly, but the multivariable

findings indicate that anchoring and loss aversion warrant particular experimental attention.

The study also contributes by separating victimization frequency from monetary loss. The monetary-loss field was descriptive: 17 participants reported positive values totaling ₹406,662, with a median of ₹2,800 and a maximum of ₹200,000. These values were not included in the inferential models and should be manually verified against the raw data before any financial-loss analysis. Keeping frequency and monetary loss separate prevents a conceptual error in which two different dimensions of harm are collapsed into one score.

4.9 Implications for fraud-prevention design

The practical implications are strongest where psychological mechanisms can be translated into interface and communication changes. First, anti-fraud systems can interrupt anchoring by requiring customers to independently confirm beneficiaries, amounts, and destination accounts rather than simply presenting attacker-supplied information. Second, high-risk transactions can incorporate deliberate pause mechanisms when unusual amounts, new beneficiaries, or high-risk channels are detected. Third, customer messages can explicitly neutralize loss-framed pressure by stating that legitimate security procedures do not require immediate disclosure of credentials or payment to “unlock” an account.

Fourth, financial institutions can separate genuine service communications from marketing or informational messages so that customers have a clearer expectation of what official alerts look like. Fifth, customer education can use scenario-based practice rather than general warnings. Training should ask users to identify the cue that is attempting to create urgency, the initial number or label that is being presented as an anchor, and the independent verification step that should occur before action.

These recommendations should not be interpreted as tested interventions in the current study. The data identify associations; they do not establish that a particular warning, pause prompt,

or interface feature will reduce fraud. Nevertheless, the convergence of the anchoring and loss-aversion results with information-processing and phishing research provides a rational basis for experimental evaluation of such interventions (Butavicius et al., 2022; Xu et al., 2025).

4.10 Limitations

Several limitations constrain interpretation. The cross-sectional design prevents temporal and causal inference. The sample of 123 was below the planned 200 and was obtained through purposive or voluntary recruitment, limiting generalizability. All key constructs were measured through self-report, creating potential common-method effects and retrospective reconstruction. Participants may also fail to recognize some encounters as fraud or may be reluctant to disclose incidents.

The victimization index combines attempted and successful incidents and therefore does not measure financial harm alone. The separate monetary-loss item is descriptive and requires verification before inferential use. The documentation mismatch between the stated 12-item victimization measure and the observable 11 frequency items plus a monetary-loss item should be corrected in future versions of the protocol. Finally, internal consistency was established, but the questionnaire has not been fully validated through factorial, convergent, discriminant, or criterion-validity analyses in this dataset.

The moderation analysis is also limited by statistical precision. Sixteen interactions were examined in a sample of 123, and none reached $p < .05$, with the smallest p approximately .053. This should be described as inconclusive rather than as definitive evidence that moderation does not exist. Future studies should pre-specify a smaller number of theoretically justified interactions and use substantially larger samples.

4.11 Future research

Future research should combine longitudinal designs with behavioral measures. Instead of asking only whether participants generally rely on

anchors or loss-framed cues, researchers can manipulate the amount, order, authority cue, urgency, and loss framing of simulated banking messages and record verification behavior. Such studies would help distinguish self-reported cognitive tendency from behavior under realistic conditions.

Larger multisite samples should also be used to examine whether the four constructs demonstrate stable factor structure and discriminant validity. Measurement invariance across demographic and banking-experience groups would help determine whether observed group differences reflect actual psychological differences or differences in how items are interpreted. Linking survey responses with verified fraud-report records, where ethically and legally feasible, would further separate attempted exposure from confirmed victimization and financial loss.

Another useful direction is intervention research. Banks and payment providers could test whether beneficiary confirmation, cooling-off periods, loss-neutral warning language, or independent verification prompts reduce unsafe actions. Randomized field or laboratory experiments would allow the mechanisms suggested by the present correlational study to be evaluated directly. Such work would move the literature from identifying vulnerability mechanisms toward testing practical countermeasures.

Finally, intervention studies should be designed around the mechanisms suggested by the present findings. Researchers could experimentally manipulate the initial amount displayed in a payment request, the presence of a threatened loss, the degree of urgency, and the availability of independent verification. Outcomes could include verification behavior, delay before action, link selection, disclosure of information, and transaction completion. Such designs would allow the field to move from correlational identification of cognitive vulnerability toward causal tests of specific countermeasures. The current findings provide a basis for prioritizing these tests, but they

should not be treated as evidence that any particular intervention is already effective.

Fifth, the distinction between self-reported tendency and observed behavior should remain central. Participants may accurately describe their general attitudes but still behave differently during an actual high-pressure interaction. Conversely, a person may report a strong tendency but never encounter a message that activates it. Combining survey measures with behavioral tasks, simulated phishing exercises, and verified incident records would therefore provide a stronger evidence base than any single method. This is particularly important because the literature shows that self-predictions and observed behavior can diverge in phishing contexts (Hassandoust et al., 2019).

Fourth, future surveys should collect incident-level context. A useful incident instrument could record the communication channel, whether the sender was known, whether urgency was present, whether a financial amount was displayed, whether credentials or a payment were requested, whether the customer independently contacted the institution, and whether the transaction was stopped. Such information would allow psychological tendencies to be linked to concrete features of fraud encounters. It would also permit researchers to distinguish exposure from compliance and compliance from financial harm.

The findings point to several measurement priorities for future research. First, attempted exposure, resisted attempts, successful unauthorized transactions, and confirmed financial loss should be measured separately. These outcomes have different meanings and may have different predictors. Second, frequency categories should be clearly defined and, where possible, complemented by exact event counts. Third, cognitive-bias scales should undergo factor analysis and validity testing before being treated as established measures across populations. Internal consistency is necessary for reliable measurement, but it does not demonstrate that a

scale has the intended factorial structure or discriminant validity.

4.13 Measurement and reporting implications

This broader framing is particularly important for psychological research on fraud because an exclusive focus on individual bias can unintentionally imply that victims are responsible for their own victimization. The more useful perspective is interactional: attackers construct cues, platforms shape choices, institutional procedures create or remove friction, and individuals interpret information under varying cognitive and emotional conditions. Human-centered prevention should therefore reduce opportunities for manipulation rather than merely instruct customers to make perfect decisions under pressure. This perspective also creates a clearer bridge between psychological theory and cybersecurity practice.

The interpretation of victimization also requires attention to exposure. Incident frequency is influenced not only by how a customer processes messages but also by how many suspicious messages the customer receives, how frequently the person transacts digitally, whether the person uses public or private networks, and whether the individual occupies social or occupational contexts in which fraud attempts are more common. These variables were not comprehensively measured. The cognitive measures therefore explain part of a broader behavioral ecology rather than the entire phenomenon. This is one reason the manuscript avoids language suggesting that cognitive bias alone determines victimization.

The study's modeling strategy also illustrates the importance of matching statistical methods to outcome structure without changing the research question after observing the results. OLS regression was retained as the primary framework because the research question concerned incremental prediction and interpretable mean differences. HC3 standard errors addressed sensitivity to heteroskedasticity. Negative-binomial regression was then used as a sensitivity analysis because the victimization index showed substantial overdispersion. The fact that

anchoring and loss aversion remained positive across these specifications provides useful robustness information, although robustness does not transform a cross-sectional association into a causal estimate.

The correlations among the four cognitive measures further support a systems interpretation. Confirmation and anchoring were strongly related, and anchoring and loss aversion were also substantially correlated. Such overlap is theoretically plausible because real-world fraud messages often combine several cues. A message can present an initial amount, claim to originate from a familiar institution, and threaten a financial or access loss in the same interaction. A respondent's score on one bias may therefore partly reflect a broader pattern of cue-sensitive processing. The regression analysis should consequently be understood as partitioning shared variance rather than identifying four independent psychological forces.

The combined pattern of results is more informative than any individual coefficient. Availability had the highest descriptive mean, yet it did not have the strongest relationship with victimization. Loss aversion had a comparatively moderate mean but the strongest bivariate and adjusted association with the outcome. Anchoring showed a similar pattern of statistical relevance. This contrast demonstrates that descriptive prominence, bivariate association, and adjusted prediction answer different questions. A construct can be common in a population without being the construct that explains the greatest amount of variation in a particular behavioral outcome.

4.12 Integrated interpretation of the findings

An additional implication concerns replication and reporting discipline. Because the present sample was smaller than planned, confidence intervals should remain central to interpretation, and future studies should be designed around a priori power calculations for both main effects and interactions. Replication should also preserve the distinction between the primary model and sensitivity analyses. Reporting only the most favorable specification can exaggerate certainty,

whereas presenting convergent and divergent estimates allows readers to evaluate robustness directly. The present manuscript therefore retains the primary HC3 regression results and reports the negative-binomial analysis as a sensitivity check. This transparent structure is preferable to treating one model as definitive when the outcome has a highly dispersed distribution.

The manuscript therefore treats the observed associations as evidence for targeted hypothesis generation. The strongest next step is not broader speculation, but direct testing of the message features that correspond to the statistical pattern: initial numerical cues, threatened losses, urgency, and opportunities for independent verification. This approach can strengthen both psychological theory and practical fraud-prevention evaluation.

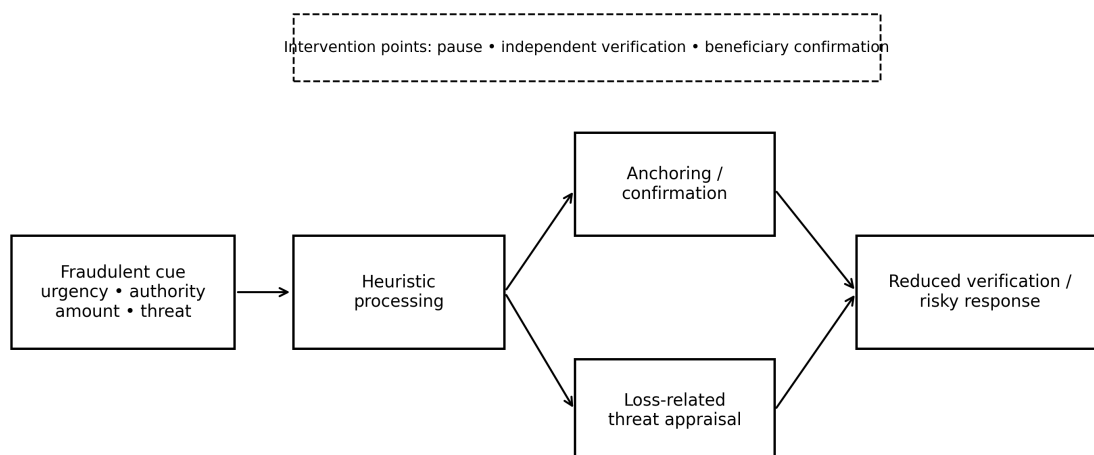
V. CONCLUSION

This study examined the relationship between four cognitive-bias constructs and self-reported digital-fraud victimization among 123 retail banking customers. The evidence indicates that cognitive-bias measures jointly added explanatory information beyond demographic controls, with anchoring and loss aversion showing the clearest adjusted positive associations. These findings were directionally reinforced by negative-binomial sensitivity analysis. Confirmation bias and availability heuristic did not demonstrate independent positive effects in the principal adjusted model.

The findings should be understood as probabilistic and contextual rather than deterministic. A cognitive tendency does not make an individual destined to become a fraud victim, and the study does not establish that any bias causes fraud. Instead, the results suggest that fraudulent communications may warrant particular attention when they use salient initial cues or threatened-loss framing. Recent analysis of phishing emails also demonstrates that multiple cognitive-bias mechanisms can be embedded in deceptive message construction (Yao et al., 2025). Human-centered fraud prevention can respond by inserting independent verification, transaction friction at appropriate points, and communication that reduces the perceived need for immediate compliance.

The study also demonstrates the importance of precise outcome construction. Attempted and successful incidents were intentionally retained within a frequency index, while monetary loss was kept separate. This makes the statistical interpretation more defensible and highlights why future fraud research should distinguish exposure, susceptibility, successful victimization, and financial harm. Larger samples, validated behavioral measures, longitudinal designs, and experimentally tested interventions are needed to determine whether the mechanisms identified here translate into measurable reductions in fraud risk.

Figure 4. *Integrated Cognitive Processing and Intervention Pathway for Digital-Fraud Prevention*



Note. The pathway integrates the study's theoretical interpretation with potential intervention points. It is a conceptual synthesis and does not represent a tested causal intervention.

REFERENCES

- Butavicius, M. A., Taib, R., & Han, S. J. (2022). Why people keep falling for phishing scams: The effects of time pressure and deception cues on the detection of phishing emails. *Computers & Security*, 123*, 102937. <https://doi.org/10.1016/j.cose.2022.102937>
- Cuadra, J. M., Mandras, Y. S., Raya, M. R., Belardo, H. A. B., Lopez, R. P., & Rodriguez, J. G. E. (2025). Experiencing investment scams in the Philippines: An interpretative phenomenological analysis of victims' financial decision-making. *Review of Behavioral Finance*, 17*(5), 769–784. <https://doi.org/10.1108/RBF-03-2025-0116>
- Hassandoust, F., Singh, H., & Williams, J. (2019). How contextualisation affects the vulnerability of individuals to phishing attempts. In *PACIS 2019 proceedings** (Paper 163). Association for Information Systems. <https://aisel.aisnet.org/pacis2019/163>
- Kahneman, D. (2011). *Thinking, fast and slow**. Farrar, Straus and Giroux.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment** (pp. 49–81). Cambridge University Press. <https://doi.org/10.1017/CBO9780511808098.004>
- Luo, X., Zhang, W., Burd, S., & Seazzu, A. (2013). Investigating phishing victimization with the heuristic-systematic model: A theoretical framework and an exploration. *Computers & Security*, 38*, 28–38. <https://doi.org/10.1016/j.cose.2012.12.003>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2*(2), 175–220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Reserve Bank of India. (2025). *Annual report 2024–25**. <https://www.rbi.org.in/Scripts/AnnualReportPublications.aspx>
- Ribeiro, L., Guedes, I. S., & Cardoso, C. S. (2024). Which factors predict susceptibility to phishing? An empirical study. *Computers & Security*, 136*, 103558. <https://doi.org/10.1016/j.cose.2023.103558>
- Spicer, M., & Nurse, J. R. C. (2026). What makes people susceptible to phishing attacks: A systematic literature review and future research agenda. *Computers & Security**, 105096. <https://doi.org/10.1016/j.cose.2026.105096>
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5*(2), 207–232. [https://doi.org/10.1016/0010-0285\(73\)90033-9](https://doi.org/10.1016/0010-0285(73)90033-9)
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185*(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Vishwanath, A., Harrison, B., & Ng, Y. J. (2018). Suspicion, cognition, and automaticity model of phishing susceptibility. *Communication Research*, 45*(8), 1146–1166. <https://doi.org/10.1177/0093650215627483>
- Vishwanath, A., Herath, T., Chen, R., Wang, J., & Rao, H. R. (2011). Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model. *Decision Support Systems*, 51*(3), 576–586. <https://doi.org/10.1016/j.dss.2011.03.002>
- Xu, F., Liu, A., & Li, X. (2025). Victimization mechanisms and countermeasures in telecom network fraud: A dual-system theoretical perspective. *Frontiers in Psychology*, 16*, 1637935. <https://doi.org/10.3389/fpsyg.2025.1637935>
- Yao, Y., Zheng, K., Wu, B., Wu, C., Gao, J., Wang, J., & Yang, M. (2025). The psychological manipulation of phishing emails: A cognitive bias approach. *Computers, Materials & Continua*, 85*(3), 4753–4776. <https://doi.org/10.32604/cmc.2025.065059>