



International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Study on Trust and Acceptance in AI-Based Financial Fraud Detection

Y.Kanaka Durga (Asst.Professor)¹, Swarna Lakshmi Pratheeka²

¹ Research guide, Department of BBA, KLEF, Vaddeswaram

Email ID: ykanakadurga@kluniversity.in

² Research Scholar, Department Of BBA, KLEF, Vaddeswaram

lakshmipratheekaswarna@gmail.com

ABSTRACT :

Payment, bank and financial institutions are increasingly adopting AI to detect fraudulent transactions, while non-transparent decision processes can prove to be a hindrance to their acceptance. This exploratory research investigates respondents' perceptions regarding four constructs namely, the significance of the trustworthy-AI principles (TRU), the relation of explainability to trust (EXP) and effectiveness of AI based detection (EFF), and utility of and acceptance of explanations (ACC). A structured questionnaire based on the five-point Likert scale was filled in by 21 respondents from a sample of 132 individuals out of a target population of 197. The quantitative data was statistically analysed in Excel with descriptive statistics, Cranach's alpha, chi-square, one-way ANOVA and linear regression analysis. Respondents perceived all the four constructs with an average (on a 5 point scale) of 3.87-4.13 indicating agreement. Further, their levels of agreement did not significantly differ according to the gender and age groups, in chi-square analysis (all Fisher's exact $p > .25$), and in one-way ANOVA analysis (all $p > .41$). In multiple regression analysis, the three predictors explained 72.3% of the variance in the acceptance of explanations ($F(3, 17) = 14.76, p < .001$), and the trustworthy-AI principles was the only significant unique predictor ($b = 0.48, p = .021$). The explainability-trust was the significant predictor, without adjusting the other predictor variables, ($b = 0.81, R = .587$). Due to the small sample size and purposive sampling, these results should be regarded as preliminary, indicating that this may be a pilot for a much larger study.

Keywords: Trustworthy AI; Explainable AI; Financial Fraud Detection; User Trust; Technology Acceptance; Likert Survey.

1. Introduction

Financial fraud is costly to banks, payment providers, and customers, and the sheer volume and intensity of digital transactions today makes manual review impractical. As a result, machine-learning models scan transactions in near real-time and tell human investigators which require further scrutiny (Bolton & Hand, 2002; Ngai et al., 2011). But accuracy is only part of the equation, and the inability to open up the reasons why models draw the conclusions they do makes it hard for investigators to verify, regulators to justify, and customers to stomach when transactions are declined.

These issues are also echoed elsewhere in the broader trustworthy AI literature. Explainability, privacy, fairness, robustness and accountability have been consistently identified as key considerations (High-Level Expert Group on AI, 2019; Jobin et al., 2019; Floridi et al., 2018) with regulation, like that in the European Union's Artificial Intelligence Act, imposing requirements on high-risk systems (Regulation (EU) 2024/1689). Effectively designing a fraud detection system is thus not about optimising only one metric, but about balancing accuracy, privacy, fairness, flexibility, explainability and robustness to adversarial attacks.

Recent technical developments in explanation methods, such as LIME and SHAP (Ribeiro et al., 2016; Lundberg & Lee, 2017), allow reasons to be assigned to each individual prediction. But do these explanations result in more, or indeed any, trust and acceptance in the people working within or impacted by fraud detection? We offer an exploratory answer to these questions using a short questionnaire.

Research questions:

RQ1. Is the degree of agreement of the participants with principles of trustworthy-AI, the link between explainability and trustworthiness, the effectiveness of AI and acceptance of explanations varies based on gender or age group (chi-square tests)?

RQ2. Are there mean differences among age groups for the construct scores (one way ANOVA)?

RQ3. How strongly is rated explanatory-trust associated with acceptance of explanations (simple linear regression)?

RQ4. Predictors of acceptance of explanations (multiple linear regression): Explainability-trust and perceived effectiveness of explanations. (a) trusting-AI principles, and (b) explainability-trust.

The balance of the paper presents related work and hypotheses (Section 2), method (Section 3), results (Section 4), and discussion of implications, limitations and future work (Sections 5-7).

2. Literature Review and Hypotheses

2.1 AI in financial fraud detection

Financial services have evolved from rule-based screening to machine learning, deep learning and anomaly detection. Shamsinejad and Banirostan (2026) survey this evolution; techniques, trends, and future directions. Kopperapu (2024) and Khan (2025) demonstrate the role of machine learning and anomaly detection for anti-fraud and risk management. Islam et al. (2025) investigate emerging AI techniques for banking fraud detection. The most recent application has been testing the real-time functioning of the techniques. Begum et al. (2025) and Oliveira (2026) look at deep learning and neural networks for real-time payment detection. Pandey et al. (2026) offer a framework that uses detection and risk forecast together. Rahmati (2025) employs an adaptive graph neural network, and Sundaravadivel (2026) offers a solution to the so-called rare-event issue of highly imbalanced transaction streams. Mwangi et al. (2026) look at the extension of the target application to adaptive compliance monitoring; and Biswas et al. (2026) suggest an agentic AI framework that detects, investigates, and mitigates fraud in real-time payment systems. The application domain is expanding to other payment and transaction fields, including Ariyibi et al. (2024), tax fraud, Ayankoya et al. (2026), small-and-medium enterprises, and Danish and Badi (2025), anti-money laundering. Cloud and cybersecurity work (Parker, 2024; Subramanian, 2025; Nowak, 2022) has introduced the fraud detection task into a wider secure framework. The literature is mostly technical, measuring success via detection capabilities. Social acceptance of these solutions, then, remains the question.

2.2 Dimensions of trustworthy AI

In the fraud detection literature we rarely treat trustworthiness as a unitary construct. Instead it appears as a number of related concerns that can be grouped into dimensions.

Transparency and explainability Aljunaid et al. (2025) introduce an explainable, federated framework for transparent banking. Sugumar (2025) mix explainable AI and secure multi-modal analytics. Prajwalasimha et al. (2025) study their approach with an eye toward trustworthy and explainable fraud detection.

Privacy and Data privacy. Federated learning is the natural solution when data is sensitive. Rahmati (2025), Gopinathan (2024), and Okunola (2025) demonstrate how models can learn between institutions, without bringing all data together. Danish and Badi (2025) connect AI-driven AML and fraud solutions with legal obligations around privacy and security in digital finance.

Security and robustness. Azmat (2026) looks at defending against adversarial attacks in fraud detection. Subramanian (2025), and Parker (2024), look at system level security. These account reliability under attack as part of their definition of trustworthiness.

Responsibility and supervision. When transactions are acted upon by autonomous systems of agentic systems (Biswas et al., 2026) or monitored by compliance (Mwangi et al., 2026), who is responsible?

The bar for authoritative frameworks (e.g., the European Commission's Ethics Guidelines for Trustworthy AI (High-Level Expert Group on AI, 2019)) is not at the end of your reference list. I would suggest inserting it and (or) something on fairness, to provide support to the axes you're choosing. Cross-check the citing.

2.3 Explainability, trust and acceptance

The logic that emerges from these sources is that one follows another explicitly and a chain implicit: explainability is needed to fight opacity, opacity is needed to motivate trust, trust is needed to support acceptance. Among the items in the list, the one explicit user perception issue that links most directly to another item is that of the paper on AI-based fraud detection and customer trust in digital banking (AI Fraud Detection & Customer Trust, 2025/2026). There is a strong argument within the explainability papers (Aljunaid et al., 2025; Sundaravadivel, 2026; Prajwalasimha et al., 2025), that transparency is justified by technical and regulatory reasons, without explicitly referring to possible empirical benefits in raising trust.

On the acceptance part, having a tech acceptance base as Davis (1989) could be very informative. In that case, you could use it to define acceptance as one intention to use or rely on AI fraud detection. Once more, this reference is missing on your reference list.

Research gap. Three gaps emerge:

1. Most of the research is technical not perceptual.
2. All types of dependability are treated in isolation, not contrasted.

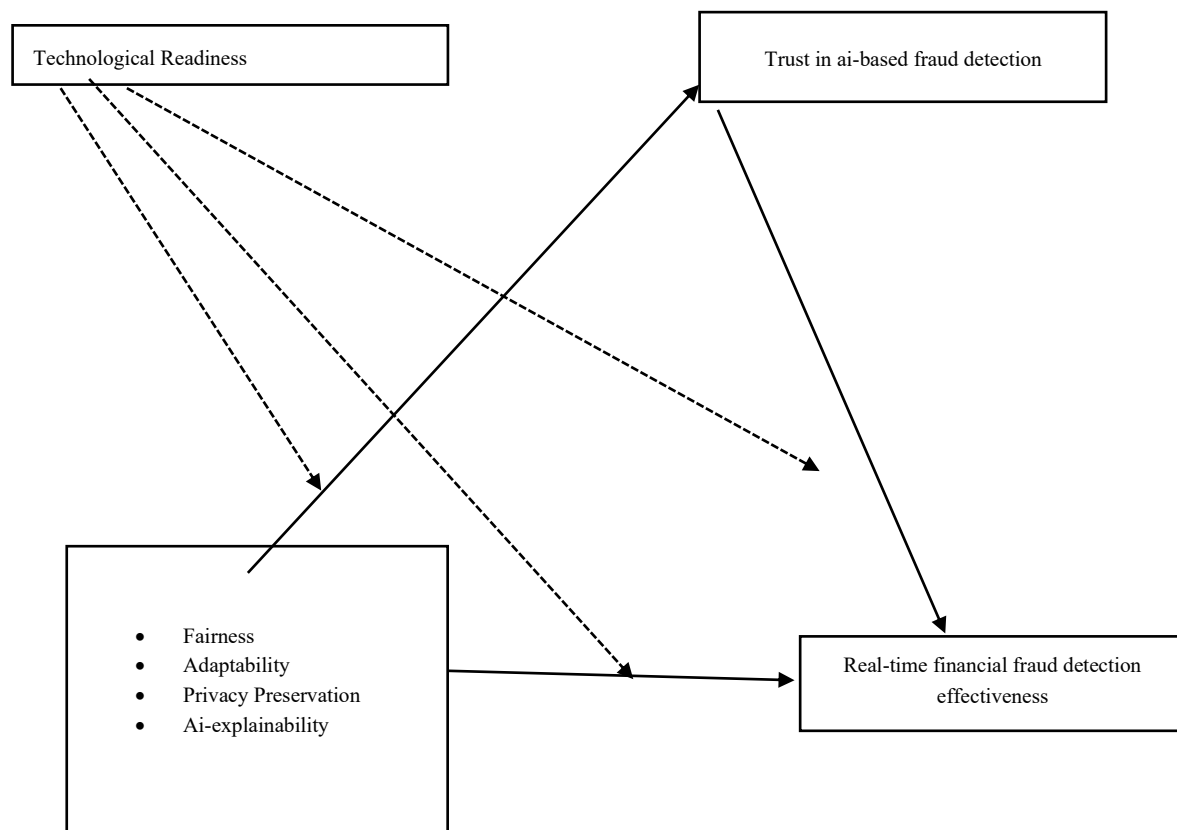
3. There is limited research identifying which dimensions are the strongest predictors of trust and acceptance.

2.4 Conceptual model and hypotheses

- ✓ *Independent variables (perceived trustworthiness dimensions):*
- ✓ *Explainability/transparency*
- ✓ *Privacy and data protection*
- ✓ *Security and robustness*
- ✓ *Accountability and oversight*

AI fraud detection acceptance Usage without any difficulty of the fraud detection interest, with ease of use Ai fraud detection adoption: actual use of it in practice

If the respondent list collected additional characteristics of respondents (such as age bracket, role, banking experience), these can also be used as grouping variables.



H01: The use of explainable AI does not impact real-time detection of financial fraud.

H02: Privacy-preserving AI does not have a meaningful impact on the efficiency of real-time financial fraud detection.

H03: Perceptions of fairness in AI based fraud detection do not significantly impact on the accuracy of financial fraud detection in real-time.

H04: Effectiveness of timely detection of financial fraud with adaptive learning.

3. Methodology

3.1 Research design

Quantitative, cross-sectional survey study. Participants responded to items related to AI-based fraud detection on a 5-point Likert scale (1= Strongly

disagree, 2= Disagree, 3= Neutral, 4= Agree, 5= Strongly agree).

3.2 Population and sample

The population being 197, the number required using the standard Yamane (1967) formula, $n = N / (1 + Ne)$, where $N = 197$, $e = .05$, is 132. The questionnaire was conducted online. 21 (10.7%) responses were received, which is 15.9% of the target sample population and 10.7% of the population size.

If the Yamane (1967) formula is inverted, $(N/n) - 1 = e^2$, then the data from these 21 respondents has a margin of error of 20.6%, so it is not generalisable to the population.

Respondent identities and their e-mail addresses were deleted prior to analysis. [Add here any approval or consent to participate by ethics committee, or consent obtained from participants, and details of how participants were recruited.]

3.3 Instrument and constructs

On this questionnaire, age group and gender were recorded and then 20 questions (Q5 to Q25 from the questionnaire numbering system) were asked, (see Appendix A). Items were placed into four different constructs based on conceptual basis.

TRU (Q5-Q9): need for explanations, privacy, adaptability, fairness and an integrated framework across accuracy, privacy, fairness, adaptability, explainability and adversarial robustness.

EXP (Q10-Q14): we believe explanations and transparency build confidence and trust and lead to better investigations.

EFF (Q15-Q18): perceived accuracy, speed and real time effectiveness of the detection based on AI.

ACC (Q20, Q22, and Q25): usefulness of "more easily understandable explanations of the functioning of" institutions and the acceptance of the detection of flaws in the AI by the instruments.

Three items (Q19, Q21 and Q23) were answered in a true/false format and were coded 5 for "true" and 1 for "false". They lack the 5 point response scale and since 20 of the 21 answered "true" they were eliminated from all constructs. Composites are averages across the construct's items and run 1-5.

3.4 Analysis procedure

All analyses were done with formulas in Excel, so results may be traced to source data. Descriptive statistics from AVERAGE, STDEV, and COUNTIF. Internal reliability from Cranach (1951) alpha derived from item and total-score variances; relationships between composites from CORREL 2-tiled T-tests; χ^2 independence tests (2 2) from observed-to-expected (CHITEST, CHIDIST), and Yates' continuity correction, and Fisher's exact test HYPGEOMDIST, because expected counts were mostly <5 (Agresti, 2013); for these composites, the age group was (high) mean > 4 (that is, Agree or above) or not, and age was collapsed to 18-23 versus 24-30.

ANOVA for composite means across four age groups (FDIST); eta squared for effect size.

Regressions from line formulas: Model 1: ACC on EXP, Model 2: ACC on TRU, EXP, and EFF; coefficients, standard errors, and VIF from inverse of correlation matrix of predictors. Significance level $= .05$; no correction was made for multiple testing. Nominal Likert composites were treated as interval scales, common in use and legitimate for multi-item scales (Norman, 2010). Results were confirmed in an independent analysis in Python.

4. Results

4.1 Sample profile

Table 1. Sample profile (N = 21)

Characteristic	n	%
Age 18-20	8	38.1%
Age 21-23	6	28.6%
Age 24-26	3	14.3%
Age 27-30	4	19.0%
Female	17	81.0%
Male	4	19.0%

Note. The age is given for the bands as indicated.

The majority of respondents were in the 18-23 age group (14 of 21; 66.7%) and 17 of 21 (81.0%) were female. The imbalance of gender and age in the sample restricts the scope of the comparisons that can be drawn later.

4.2 Descriptive statistics and reliability

Figure 1 presents the item means. Means were between 3.76 (Q15, Q16 and Q17) and 4.33 (Q10, "clear explanations increase confidence") and the proportions of respondents who agreed or strongly agreed varied from 61.9% (Q16) to 85.7% (Q10). Results at the construct level can be seen in Table 2.

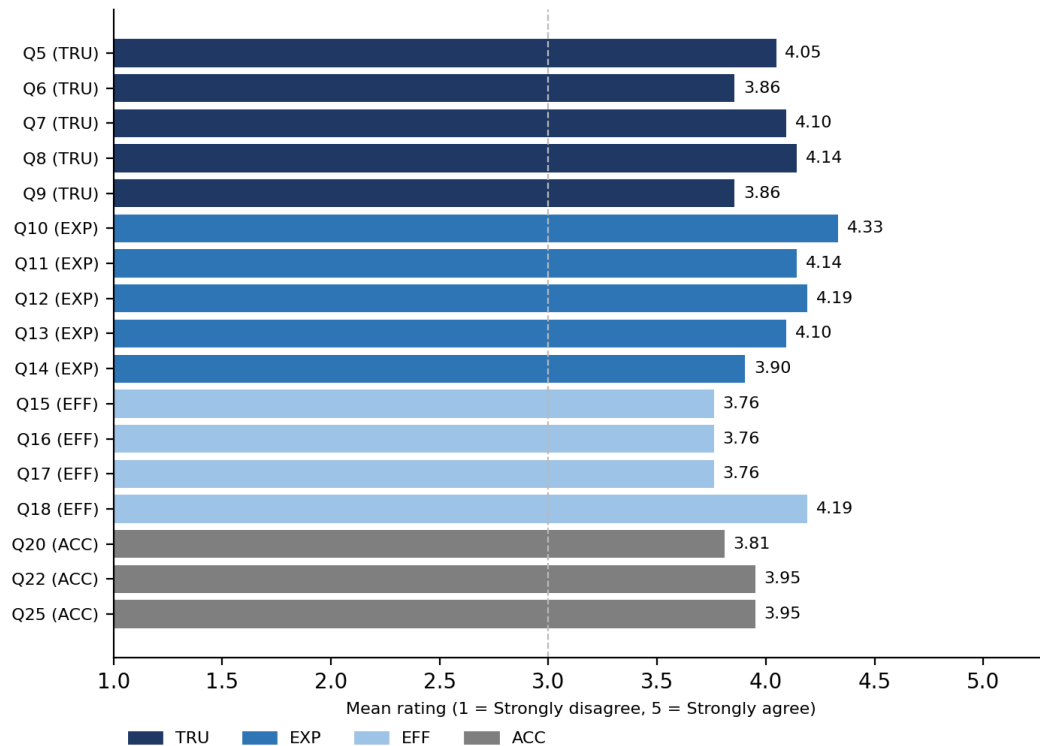


Figure 1. Mean rating for each Likert item by construct (Q19, Q21 and Q23 excluded).

Table 2. Descriptive statistics and internal consistency of constructs

Construct	Items	Mean	SD	Min	Max	Cronbach's α
TRU – Trustworthy-AI principles	Q5-Q9	4.00	0.87	2.20	5.00	.843
EXP – Explainability-trust link	Q10-Q14	4.13	0.73	2.80	5.00	.771
EFF – Perceived effectiveness	Q15-Q18	3.87	0.71	2.50	5.00	.624
ACC – Explanation utility & acceptance	Q20, Q22, Q25	3.90	0.77	2.33	5.00	.750

Note. Scales were on a 1 to 5 Likert-type scale. N=Glick et al.; b=20 communities (see previous note); c=30 communities (see previous note). Alpha at 5 is total alpha. Alphas of .70 or greater are typically considered appropriate (Nunnally & Bernstein, 1994).

Participants also responded favorably to each construct, with average ratings above the midpoint (see Table 2). The mean for EXP was the largest, at 4.13 (SD = 0.73), and the lowest for EFF was 3.87 (SD = 0.71). Reliability was excellent for TRU ($\alpha = .843$), and satisfactory for EXP ($\alpha = .771$) and ACC ($\alpha = .750$). Reliability for the EFF scale was below the .70 cut-off guideline ($\alpha = .624$), so is interpreted with care.

4.3 Correlations between constructs

Table 3. Pearson correlations between composite scores

	TRU	EXP	EFF	ACC
TRU	1			
EXP	.796***	1		
EFF	.535*	.540*	1	
ACC	.824***	.766***	.560**	1

Note. N = 21. P <.05, p <.01, p <.001 (2-tailed).

The correlations were all significant and positive. TRU, EXP and ACC correlated highly with each other ($r=.796$ for TRU-EXP, $r=.824$ for TRU-ACC and $r=.766$ for EXP-ACC). EFF also correlated modestly with each of the other three constructs (r ranging from .535 to .560).

4.4 Chi-square tests of independence

Table 4. Chi-square tests of independence (DF = 1)

Test	High / Not high (row 1)	High / Not high (row 2)	χ^2	p	ϕ	Fisher's p
Gender $\times$ TRU level	Female: 12 / 5	Male: 1 / 3	2.85	.091	.369	.253
Gender $\times$ ACC level	Female: 12 / 5	Male: 2 / 2	0.62	.432	.171	.574
Age group $\times$ EXP level	18-23: 10 / 4	24-30: 4 / 3	0.43	.513	.143	.638
Age group $\times$ EFF level	18-23: 8 / 6	24-30: 3 / 4	0.38	.537	.135	.659

Note. "High" = Composite mean 4.0. Rows are Female/Male for gender tests and 18-23/24-30 for age tests. Smallest expected counts were 1.52, 1.33, 2.33 and 3.33, which are less than 5, so the Fisher's exact test result is given. = phi coefficient.

Neither of the 4 tests reaches the .05 level. The only trend is for gender and TRU ($(1) = 2.85$, $p = .091$, $\phi = .369$): 12 of 17 women but only 1 of 4 men is "High". It is a medium-to-large effect size but it is not statistically reliable (Fisher's $p = .253$) and no conclusion can be reached because there are only 4 men. Gender is not related to level of ACC ($p = .432$; Fisher's $p = .574$); age group 18-23 vs 24-30 is not related to level of EXP ($p = .513$; Fisher's $p = .638$) or level of EFF ($p = .537$; Fisher's $p = .659$). In each case expected counts are so small the tests have very low power.

4.5 One-way ANOVA by age group

Table 5. Construct means by age group and one-way ANOVA results

Construct	18–20 (n = 8)	21–23 (n = 6)	24–26 (n = 3)	27–30 (n = 4)	F(3, 17)	p	η^2
TRU	4.13 (0.72)	4.10 (0.71)	3.93 (1.01)	3.65 (1.45)	0.27	.848	.045
EXP	4.20 (0.76)	4.20 (0.64)	4.00 (0.92)	4.00 (0.94)	0.10	.958	.018
EFF	3.72 (0.60)	4.29 (0.64)	3.67 (0.58)	3.69 (1.03)	1.01	.412	.151
ACC	3.96 (0.70)	4.00 (0.99)	3.78 (0.38)	3.75 (1.00)	0.11	.954	.019

Note. Cells are means (SD). F critical ($\alpha = .05$; $df = 3, 17$) = 3.20. = eta squared.

There was no significant differences for any construct across age groups: TRU, $F(3, 17) = 0.27$, $p = .848$, EXP, $F = 0.10$, $p = .958$, EFF, $F = 1.01$, $p = .412$, ACC, $F = 0.11$, $p = .954$. All F scores were well below the critical value of 3.20. For all constructs, the amount of variance due to the independent variable (eta squared) was highest for EFF ($= .151$), with 21-23 year olds having the highest responses ($M = 4.29$), but it was not statistically significant. The aged 24-26 and 27-30 groups only had 3 and 4 respondents so comparisons between these groups are both underpowered and unstable.

4.6 Linear regression

Model 1 (simple regression): ACC regressed on EXP. The regression was significant, $F(1, 19) = 27.02$, $p < .001$, $R = .587$ (.565 Adjusted R). For every one point increase in EXP, ACC increased by .81 points, $b = 0.81$, $SE = .16$, $t(19) = 5.20$, $p < .001$, 95% CI [0.48, 1.13].

The line of best fit is: $ACC = 0.57 + 0.81 EXP$.

Figure 2: The Regression Line for H1. Thus, H1 is supported at the bivariate level.

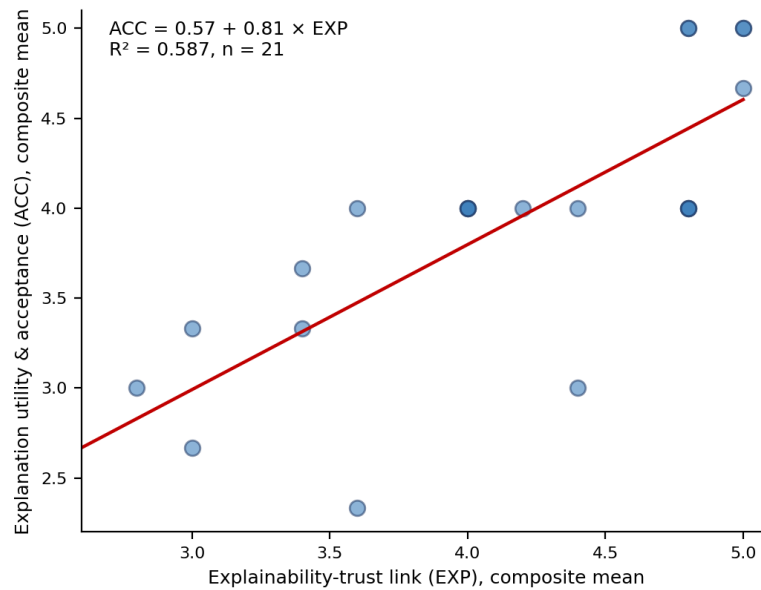


Figure 2. Relationship between EXP and ACC composite scores with fitted regression line (several respondents share identical scores, so points overlap).

Table 6. Multiple regression predicting acceptance of explanations (ACC), Model 2

Predictor	b	SE	t	p	95% CI	β	VIF
Intercept	0.31	0.63	0.49	.631	[-1.03, 1.64]	–	–
TRU	0.48	0.19	2.54	.021	[0.08, 0.88]	0.55	2.85
EXP	0.28	0.23	1.21	.241	[-0.20, 0.76]	0.26	2.87
EFF	0.14	0.17	0.80	.432	[-0.22, 0.49]	0.12	1.48

Note. N = 21. R = .723, adjusted R = .674, F(3, 17) = 14.76, p < .001, standard error of estimate = 0.44. = standardised coefficient; VIF = variance inflation factor.

Model 2 (multiple regression) tested H2 and H3. This was significant and accounted for 72.3% of the variation in ACC (adjusted R = .674), F(3, 17) = 14.76, p < .001. The only significant predictor at the multivariate level was TRU (b = 0.48, SE = 0.19, t(17) = 2.54, p = .021, β = 0.55). EXP (b = 0.28, p = .241) and EFF (b = 0.14, p = .432) were not significant once TRU was included in the regression equation. Variance Inflation Factor (VIF) for the three predictors was below 3 (2.85, 2.87 and 1.48), therefore collinearity was not a problem, but the fact that the correlation between TRU and EXP was so high (r = .796) indicates that the two may account for much of the same variance. H2 is thus supported, and H1 and H3 are not supported at the multivariate level.

4.7 Summary of hypothesis tests

Table 7. Summary of hypotheses

Hypothesis	Evidence	Outcome
H1: EXP → ACC (positive)	Simple regression: b = 0.81, p < .001, R ² = .587; in multiple regression b = 0.28, p = .241	Supported alone; not unique
H2: TRU → ACC (positive)	Multiple regression: b = 0.48, p = .021, β = 0.55	Supported
H3: EFF → ACC (positive)	Multiple regression: b = 0.14, p = .432	Not supported

Note. Outcomes are preliminary because of the small sample (N = 21).

5. Discussion

The respondents were broadly in favour of these principles of trustworthy, explainable AI for fraud detection. In every case the mean scores for the statements were above 3.8 and the strongest single statement was for the items relating to the advantages of clear explanations (Q10, M = 4.33). The

mean scores for the items relating to the advantages of AI itself (Q15-17, $M = 3.76$) were all lower, indicating that while respondents ascribe to the principles, they are not necessarily as confident in the accuracy and speed of the systems. This could suggest that explanations and safeguards are valued because performance claims are not enough, but this cannot be tested here.

Between-group differences according to chi-square and ANOVA are not reliable. However, this lack of significance should not be taken as that of nonexistence: one can determine from the tests used that there were very few men and several respondents in the 24-30 age grouping in a sample of 21, which was not sensitive to anything but very large effects. A couple of the effect sizes are large enough to conceivably demonstrate real differences in a bigger sample.

Regression results are the most interpretable. Explanation usefulness-acceptance explained 59% of the variation in explanation acceptance when perceived explainability-trust was the only predictor, as might be expected if people who trust explanations for reasons of explanation are also to find explanations useful. The constructs explainer-effectiveness and explanation, when TRU was added, explained no more than explainability-trust on its own and explainer-effectiveness was replaced by TRU. It is reasonable to hypothesize that people who find all trustworthy-AI principles useful (privacy, fairness, explanation, adaptability) are also people who find explanations useful for practical purposes, so the attitude toward explanation that explains general trust in explanations might also be a general attitude toward responsible AI which explains the response to the items. Alternatively, the overlapping item content (Q5 of TRU about explanations, ACC items about explanations) could make the constructs overlap very strongly. These two accounts cannot be distinguished with the data. The non-significance of the explainer-effectiveness coefficient implies that belief in AI accuracy and speed adds little to explanation acceptance beyond trust attitudes.

6. Practical Implications

While still preliminary, our findings lead to three recommendations to think about for any institution building or implementing AI-based fraud detection. 1. Explanation features should be framed within a general and broad trustworthy-AI message that incorporates explanation with privacy and fairness, since the strongest correlation with explanation acceptance was endorsement of all of those principles together. 2. Investigators and decision makers seem to value explanations for both understanding a contributing factor to the alert and improving a decision, so explanatory output should be designed to fit investigators' workflows (e.g., explaining the factors behind an alert). 3. Explanation of detection quality, including false-alert rates, is still required, since respondents rated AI effectiveness lower than AI principles.

7. Limitations and Future Research

Limitations (1) The sample size was small and a convenience sample; 21 of 132 planned participants were recruited, limiting generalisability and resulting in an expected margin of error of 21%. (2) Most participants were young women and only four were men. (3) With seven observations per predictor, the regression was underpowered; based on Green (1991), a sample size of around 74 or more is suggested for testing overall fit with three predictors, and over 100 for individual predictors, and the coefficient estimates here are unstable. (4) All measures are self-report at one time, so no causal inferences are justified. (5) The scale measuring EFF had a modest scale reliability for ACC ($\alpha = .62$) and three true/false items, whose responses were recoded, were dropped from ACC resulting in three ACC items. (6) Multiple constructs shared some item content and chi-square and ANOVA testing was not corrected for multiple testing. (7) Likert responses were treated as interval data.

Future research should be gathered to full data for the proposed sample, show the all items with 5 points scale, and incorporate respondents with the direct experience of fraud-investigation. And the larger sample provide an opportunity to test the construct measurement (such as Confirmatory Factor Analysis), structural equation modelling the relationships among constructs, and test differences across stake holder groups including investigators, customers, and compliance staff. Experimental design showing or not showing explanations would test whether explanations actually affect trust and decision quality.

8. Conclusion

The exploratory nature of this survey suggested some encouraging findings on trusts and understandableness (trustworthy-AI attitudes) in financial fraud detection, no differences by gender or age group and a surprisingly high regression fit where the support for trustworthy-AI principles was the main predictor for explainability ($R = .723$). Only 21 responses were received so the data are preliminary. The findings will be replicated with the combined survey in the future with a more traditional Likert scale version.

Declarations

Ethics and consent: [Insert statement.] Funding: [Insert or none.] Conflicting interests: [Insert or none.] Availability of data and material: The analysis workbook (TrustworthyAIFraudDetectionAnalysis.xlsx) holds the item responses anonymised and all formulas are set out in this paper.

REFERENCES

- Aljunaid, S. K., Almheiri, S. J., Dawood, H., & Khan, M. A. (2025). Secure and transparent banking: Explainable AI-driven federated learning model for financial fraud detection. *Journal of Risk and Financial Management*, 18(4), 179.

Full link: <https://doi.org/10.3390/jrfm18040179>

2. AI fraud detection & customer trust: A study on AI-based fraud detection and customer trust in digital banking. (2025/2026).

Full link: <https://scholar.google.com/scholar?q=%22AI+Fraud+Detection%22+%22Customer+Trust%22+digital+banking>

3. Rahmati, M. (2025). Real-time financial fraud detection using adaptive graph neural networks and federated learning. *International Journal of Management and Data Analytics*, 5(1), 98–110.

Full link: <https://ijmada.com/index.php/ijmada/article/view/77>

4. Shamsinejad, E., & Baniroostam, H. (2026). A comprehensive survey on AI-driven financial fraud detection: Trends, techniques, and future directions. *FinTech and Sustainable Innovation*.

Full link: <https://ojs.bonviewpress.com/index.php/FSI/article/view/8230>

5. Pandey, C. P., Upadhyay, H., Kale, A., Joshi, P., Katta, B. S., & Kumar, R. (2026). AI-driven fraud detection and risk forecasting framework for real-time financial transactions. *Scientific Culture*, 12(1.1), 3425–3431.

Full link : <https://scholar.google.com/scholar?q=%22AI-Driven+Fraud+Detection+and+Risk+Forecasting+Framework%22>

6. Sugumar, R. (2025). Explainable AI-driven secure multi-modal analytics for financial fraud detection and cyber-enabled pharmaceutical network analysis. *International Journal of Advanced Research in Computer Science & Technology*, 8(6).

Full link: <https://www.ijarcst.org/index.php/ijarcst/article/view/354>

7. Sundaravadivel, P. (2026). Federated deep learning and explainable AI for real-time credit card fraud detection in highly imbalanced transaction streams. *Discover Artificial Intelligence*, 6, 858.

Full link: <https://doi.org/10.1007/s44163-026-01449-z>

8. Khan, M. A. (2025). AI-based anomaly detection in financial fraud prevention systems. *Multidisciplinary Research in Computing Information Systems*, 5(12).

Full link: <https://scholar.google.com/scholar?q=%22AI-Based+Anomaly+Detection+in+Financial+Fraud+Prevention+Systems%22>

9. Islam, I., Chowdhury, S. A., Hoque, A., & Hasan, M. M. (2025). The future of banking fraud detection: Emerging AI technologies and trends. *Well Testing*, 34(S3), 99–117.

Full link: <https://scholar.google.com/scholar?q=%22The+Future+of+Banking+Fraud+Detection%22>

10. Islam, I., Chowdhury, S. A., Hoque, A., & Hasan, M. M. (2025). The future of banking fraud detection: Emerging AI technologies and trends. *Well Testing*, 34(S3), 99–117. [Duplicate of Paper 9].

Full link: <https://scholar.google.com/scholar?q=%22The+Future+of+Banking+Fraud+Detection%22>

11. Danish, A., & Badi, S. (2025). AI-powered AML and fraud detection: Protecting financial institutions while enhancing data privacy and security in digital finance.

Full link: <https://scholar.google.com/scholar?q=%22AI-Powered+AML+and+Fraud+Detection%22>

12. Kopperapu, R. (2024). Harnessing AI and machine learning for enhanced fraud detection and risk management in financial services. *International Research Journal of Economics and Management Studies*, 3(12), 109–114.

Full link: <https://irjems.org/irjems-v3i12p113.html>

13. Sundaravadivel, P. (2026). Federated deep learning and explainable AI for real-time credit card fraud detection in highly imbalanced transaction

streams. Discover Artificial Intelligence, 6, 858. [Duplicate of Paper 7].

Full link: <https://doi.org/10.1007/s44163-026-01449-z>

14. Begum, S., Hassaan, A., Hussain, A., Jamshaid, M. M., Akbar, Z., Niaz, S., & Jobiullah, M. I. (2025). AI-driven fraud detection in real-time financial transactions: A deep learning approach.

Full link: <https://scholar.google.com/scholar?q=%22AI-Driven+Fraud+Detection+in+Real-Time+Financial+Transactions%22>

15. Prajwalasimha, S. N., Saini, D. K. J. B., Shelke, N., Pimpalkar, A., Kumar, G. H., & Pooja Shree, H. R. (2025). Trustworthy and interpretable AI for robust fraud detection in financial transactions. In 2025 6th International Conference for Emerging Technology (INCET). IEEE.

Full link: <https://scholar.google.com/scholar?q=%22Trustworthy+and+Interpretable+AI+for+Robust+Fraud+Detection%22>

16. Gopinathan, V. R. (2024). Cyber-resilient digital banking analytics using AI-driven federated machine learning on AWS. International Journal of Engineering & Extended Technologies Research, 6(4), 8419–8426.

Full link: <https://www.ijeetr.com/index.php/ijeetr/article/download/181/167/328>

17. Ayankoya, M. B., Onyemakonor, E. O., & Isibor, F. O. (2026). Developing AI-driven predictive analytics for enhancing financial forecasting, fraud detection, and operational resilience in U.S. small and medium enterprises (SMEs). International Journal of Management and Organizational Research, 5(4), 15–21.

Full link: <https://scholar.google.com/scholar?q=%22Developing+AI-Driven+Predictive+Analytics%22+fraud+detection>

18. Parker, J. O. L. (2024). AI-enhanced cloud security architecture: Machine learning-driven fraud detection for preventing business financial loss. International Journal of Advanced Research in Computer Science & Technology, 7(1).

Full link: <https://ijarct.org/index.php/ijarct/article/download/302/296/602>

19. Azmat, H. (2026). Defense strategies against adversarial attacks in AI-based financial fraud detection systems.

Full link: <https://scholar.google.com/scholar?q=%22Defense+Strategies+Against+Adversarial+Attacks%22+financial+fraud>

20. Subramanian, R. S. (2025). AI-driven cybersecurity software for fraud and network intrusion detection in financial big data systems. International Journal of Advanced Research in Computer Science & Technology, 8(6), 13218–13224.

Full link: <https://ijarct.org/index.php/ijarct/article/view/326>

21. Nowak, A. M. (2022). Secure AI-enabled cloud platforms for healthcare image analysis and financial fraud detection across web applications and 5G networks. International Journal of Engineering & Extended Technologies Research.

Full link: <https://scholar.google.com/scholar?q=%22Secure+AI-Enabled+Cloud+Platforms%22+financial+fraud>

22. Mwangi, D. K., Rodriguez, M. S., & Adebayo, A. F. (2026). Adaptive and real-time AI-driven financial compliance and fraud monitoring.

Full link: <https://scholar.google.com/scholar?q=%22Adaptive+and+Real-Time+AI-Driven+Financial+Compliance+and+Fraud+Monitoring%22>

23. Oliveira, D. N. O. (2026). Neural networks for real-time financial fraud detection. Journal International Review of Research Studies, 1(3), 1–49.

Full link: <https://doi.org/10.66104/xkad2306>

24. Biswas, B., Faysal, S. M., Das, S., Hanif, A., Akter, T., Rashed, R. A. M., & Shamarukh, S. (2026). A framework for agentic AI-driven financial fraud detection, investigation, and risk mitigation in real-time payment ecosystems. International Journal of Computer Information Systems and Industrial Management Applications, 18(3s), 1029–1040.

Full link: <https://scholar.google.com/scholar?q=%22A+Framework+for+Agentic+AI-Driven+Financial+Fraud+Detection%22>

25. Okunola, C. O. (2025). Federated learning for privacy-preserving AI. *Journal of Cyber-Physical Security and Robotics*, 1(2), 1–8.

Full link: <https://jocpsr.com/index.php/journal/article/view/15>

26. Ariyibi, K. O., Bello, O. F., Ekundayo, T. F., Oladepo, O. I., Wada, I. U., & Makinde, E. O. (2024). Leveraging artificial intelligence for enhanced tax fraud detection in modern fiscal systems. *GSC Advanced Research and Reviews*, 21(2), 129–137.

Full link: <https://doi.org/10.30574/gscarr.2024.21.2.0415>