ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

HUMAN-IN-THE-LOOP VS. LIGHTS-OUT AUTOMATION

** Mayuresh Vijay Pabarekar, ** Vedant Rambahadur Singh,*** Atharva Dinesh Pagade & **** Arman Sameer Momin*

* *Students*

Abstract:

As the global financial landscape is transitioning from a traditional ecosystem to digitally based transactions rapidly, AI models for scam detection have turned into an indispensable pillar for proactive defence. The opaqueness of automated decision-making has however caused the Human Trust conundrum that cannot be addressed by algorithmic models alone. To find out the levels of trust among the users, this paper involves comparison of consumer trust in human-in-the-loop systems versus that of completely independent systems.

The mean trust score of AI systems of the survey was calculated using inferential statistics and a 5-point Likert scale, the measure of trust, and provided a mean of 3.04 and a standard deviation of 0.30, compared to the average score of a human-led systems at 2.97 and a standard deviation of 0.32. The t-test of paired samples gave a value of 0.5984, that is, there is no considerable difference between the levels of trust between the two interventions. Moreover, ANOVA tests confirmed that trust level do not differ based on age group. Chi-square provided a p-value of 0.0529, that is, there is a strong marginal tendency preferring age- based preference, although 53.45% of the participants chose instantaneous AI blocking as opposed to human verification (46.55%). The result of these findings is a situation where trust is balanced, where users perceive that the respective systems are equally reliable, despite the possibility that functional preferences for speed may be directed towards a gradual shift to automated solutions.

Introduction:

The global finance industry is undergoing a sea change due to Artificial Intelligence (AI) and Machine Learning (ML). As digital transactions become a part of everyday life, the quantity of data has long surpassed what humans can handle. As a result, AI-based fraud detection, which used to be an innovation, has become a "prerequisite", giving security in real-time without any human intervention.

In spite of AI, it has the advantage of being fast and better at detecting fraud than humans, but it still faces real-world difficulties when actually putting it to work. In the domain of cyber-financial security, an interesting paradox exists: while systems provide speed, their lack of transparency creates user apprehension toward the automation of decision-making, particularly to "False Positive"(where an AI incorrectly blocks a safe transaction). This creates new variable called "Human Trust," which is non-computable and cannot be replaced by code alone.

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

Although much research exists on the veracity of AI models, there is very little quantitative evidence regarding user sentiment in the face of "Lights-out Intervention". What makes customers feel more secure? An automated message from AI or a manual verification Protocol?

The research paper's aim is to find out the trust levels of the users for AI-based automated systems (like those used by HDFC Bank or ICICI Bank) versus human-led verification. By leveraging Inferential Statistical methods, in this study, we investigate if a significant deviation exists on how modern humans react to cognitive appraisal of automated vs. manual interventions.

Statement of problem:

The research paper is conducted to calculate the trust score of humans. Does the consumer trust the AI detection module for speed-based blocking or a human verification call for personal assurance?

The study concerns which method the sample believes is more likely to cause a "false alarm": is it the AI- based system or the human-verification system? Does the result vary based on age group?

Significance of the study:

There has been a lot of research regarding which framework is the most efficient: automated detection systems or human integration with these systems when it comes to fraud detection and financial scams. The only thing these studies seem to miss is what the human perception of trust is while dealing with the respective systems. This is where our study comes in, where we have investigated how trustworthy and reliable people consider these frameworks. In this era of uncertainty, where people are not certain about what the future of AI will be, it is very important to find out their socio-psychological response while dealing with these frameworks. As the stakes are very high in the financial sector while facing any fraudulent activity, the trustworthiness and reliability of the framework become the most important factors**.**

Limitations:

Similar to any research study this one also has certain limitations that should be taken into account when interpreting the results.

The most apparent restriction is the age profile of the respondents. Of a total of 117 valid responses, 76 percent fell in the 18-25 years old bracket. This implies that the research is largely based on the perceptions of young people that tend to be more at ease with technology. The elderly participants did not feature equally particularly those over 60 years (only 2 respondents). Due to this disparity, the findings might not be a complete reflection of the views of senior banking clients.

A problem with a small sample size among certain age groups is also an issue. The statistical tests of differences in age is not a factor powerful enough to achieve smaller differences, as the size of participants was less in older categories. The Chi-Square result is close to significant change (0.0529), which implies that larger and more equated sample would have revealed definite differences between generations.

The research relies on self-reported answers gathered through the use of a Likert scale. These reactions indicate the way participants believe they would respond, whereas actual action in a real situation of a fraud may be different. Individuals can react differently to a genuine economic danger.

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

Moreover, the survey was conducted within a limited time frame (between February 5 and 6, 2026) and a geographical location was also used. The banking system and use of technology can be distrusted in different locations, and therefore, the results would not be universal.

Lastly, the study made a general comparison of AI and human systems without making specific comparisons between various types of AI models or the degree of human expertise which have differences in their impact on trust levels

Objectives:

1. To measure the mean trust score between AI-based systems and Human-operated systems which is measured on a 1-5 Likert scale.
2. To find out if any statistical difference is present between AI systems and Human systems, measured using a paired samples t-test.
3. To understand the impact of age on consumer trust for AI and human system using one-way ANOVA test.
4. To analyse consumer choice between immediate AI block for speed and human verification call for assurances using an independent chi-square test.
5. To spot trends, patterns or correlations between age groups and behavioural choices to give practical insights for financial sector.

Hypothesis of the study:

Ho: There is no significant differences in the mean trust score between Ai fraud detection system and human fraud detection system

H1: There is significant differences in the mean trust score between Ai fraud detection system and human fraud detection system

Review of literature:

1. **Transition towards a new approach:** The global finance industry is undergoing a significant shift in its operations. While it was once a popular trend of automation-centered industries, it is now transitioning into a collaborative intelligence industry. As the technological landscape keeps on maturing, the trend is shifting from completely automated artificial systems to human intuition integrated with machines, which is argued by Ali et al. (2024) in their research. This shift is essential as cyber and financial threats are escalating as well as evolving at a rapid pace, which demands a human-centric approach integrated with advanced technological capability for prevention and damage control.
2. **The Standard Practice:** To accomplish the goal of integrating human intuition and reasoning with automatic algorithmic-based models, for achieving desired results and precision decision-making in the finance and banking sector, it requires a sound technological and hardware foundation which would be capable of managing huge volumes of transactions and queries in real-time. According to the present trend, the complexity and volume of these transactions and queries is growing at a rapid pace on a year-on-year basis, which certainly makes it unfeasible as well as non-reliable to rely upon manual methods for detecting fraudulent patterns and financial frauds. As a result, implementing artificial models has become the standard

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

practices within the industry because of its instant detection of fraudulent patterns within the transactions and real-time proactive defence, which cannot be achieved through conventional ways as noted by Venu (2025).

3. **The Socio-psychological Factor:** Despite the technical prowess and most efficient security protocols, these automated models are not able to cope with the socio-psychological expectations of a human-centric society, which mainly focuses on trust and transparency. In the present and upcoming times, speed and accuracy are not the only evaluation metrics that matter the most in the banking and finance industry, but a transparent justification for why the decision was taken. Research by McNally and Bastos (2025) argues a similar point that the effectiveness of automated systems completely depends upon the transparent and trustworthy nature of the system. Ultimately, when automated systems function without a clear justification, it risks complete breakdown of foundational trust in banking and financial institutions.
4. **The Collaborative Framework:** To overcome a situation where trust and transparency of financial institutions backfire, the industry is shifting towards a safer and better framework that advocates collaboration between human and artificial systems, known as the "hybrid approach". In this framework, the heavy computational work, such as identifying patterns from huge volumes of datasets and automated detection and alerts are handled by artificial systems, while the tasks that need human intervention and reasoning for final decision-making are done by human analysts based on the conclusions of tasks filtered by artificial systems. Research by Yeo et al. (2025) also stresses a similar collaborative framework, the "Human-in-the-Loop" (HITL) architecture, emphasizing that this is the most effective way to ensure transparency and customer trust when AI is deployed for high-risk financial tasks. This collaborative framework, which works upon a hybrid approach, ensures that the hybrid synergy is technically as well as socially reliable.

Research Methodology:

The main motive of the study was to investigate the trust levels of the consumers on completely automated systems versus that of human intuition integrated with machines. A systematic approach was followed while researching. This section elaborates on the tools and methods which were used to gather the data from the target group.

Research Design:

The study follows a quantitative research design to collect structured statistical data and interpret the patterns observed within present consumer data to evaluate user preference and level of trust among the users while dealing with completely self-independent models and human-collaborated systems.

Data Sampling and Participants:

In our study, we have only included primary data which consist of more than 150 respondents. The participants were chosen based on their priority for making use of digital payment versus traditional payment methods. Consumers preferring digital payment systems were only included. To gather responses from various categories of consumers we have made use of convenience sampling.

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

Data-Collection Tool:

An online survey was conducted using Google Form as a tool for primary data gathering which was structured into multiple parts into specific required section as per our research objectives.

The following sections were in the survey:

Demographics and Background: In this section, we have gathered age demographics as metrics for comparison as well as their preferred banking method. Respondents have two options: traditional banking (physical branches, net banking) along with digital-only fintech (Gpay, Revolut, Phonepe, Neo-banks). Their encounters with fraudulent transactions and security alerts were also considered.

Trust Equilibrium: This section includes a Likert scale which does a sentiment analysis of respondents by recording their trust score in AI and humans on a scale of 1-5. We have also recorded their preference if any suspicious and fraudulent transaction occurs.

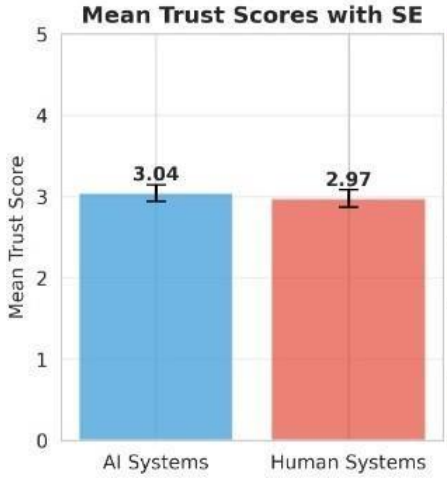
Data Analysis and Interpretation:

1. ***Descriptive Statistics***

The overall aim currently is to determine consumer confidence in systems that are driven by people and also those driven by artificial intelligence. The trust score is established using a Likert scale, and 1 is low trust and 5 high trust.

The Descriptive Statistics contains:

1. The mean score of trust in the AI-based system is 3.04 and the standard deviation is 1.09.
2. The average score of the human-operated system regarding trust is 2.97 with 1.15 as standard deviation.



Metric	AI Systems	Human System
Mean	3.04	2.97
Median	3.0	3.0
Std Dev	1.09	1.15
Min	1	1
Max	5	5

It demonstrates that the trust of people is neutral (3.0), and that of artificial intelligence polarized between low (35%), and high (34.2%).

2. ***Hypothesis Testing in Primary Paired Samples T-Test***

The paired-samples t-test is applied to discover is there any difference in the trust score of AI-based systems and human-operated systems equals 0.07 and this difference is statistically important.

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

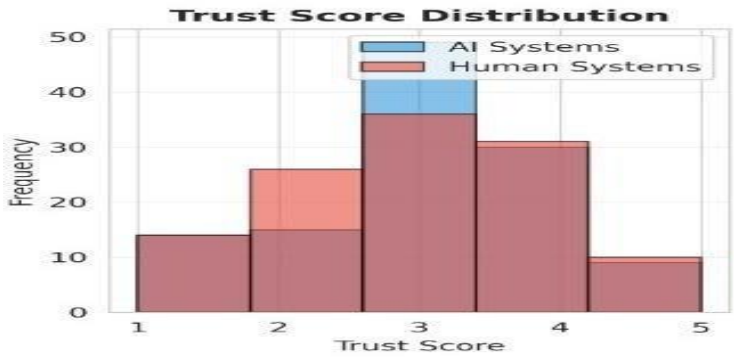
OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

Null Hypothesis (H0): There is no subtle difference in the average score of human-operated systems trust and the mean score of AI-based systems trust.

Alternative Hypothesis (H1): There is a subtle difference in the average of the human-operated systems trust score and the mean AI-based systems trust score.



Since the test shows that the t-statistic was 0.5282 with a p-value of 0.5984 which is more than alpha of 0.05 therefore, null hypothesis is accepted. That indicates there is no subtle difference in the confidence scores of the AI-based systems and the human-operated systems in the financial and banking industry.

3.*Demographic Analysis (ANOVA)*

The question that was answered by the one-way ANOVA test was whether or not the trust score varies among a number of age groups.

Age and AI trust:

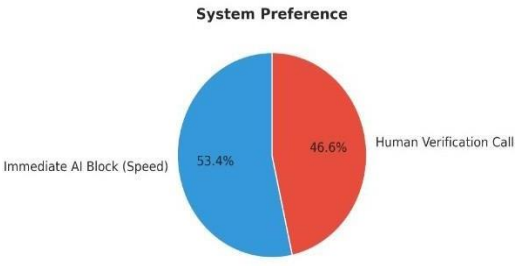
F (3,113) = 0.175; p-value = 0.913.

F (3,113) = 0.652, p-value = 0.584 by Age in human trust.

The p-values are greater than 0.05; therefore, this means that the trust score does not have a significant difference among a number of age groups.

4.*Preference, Chi-Square Test*

To analyze consumer preference between instant AI block of speed and human verification call of assurances, Chi-square test was implemented Selection: 53.45 selected the AI block; 46.55 selected human verification.



The test gives a p-value of 0.0529. It indicates a peripheral likelihood with a greater proportion of the sample disclosing youthful inclinations towards the AI block and seasoned inclinations towards human validation.

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a)

Jan – Feb 2026

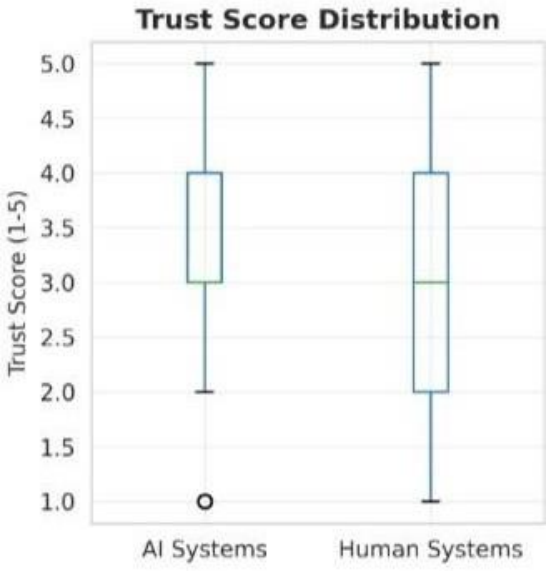
Original Research Article

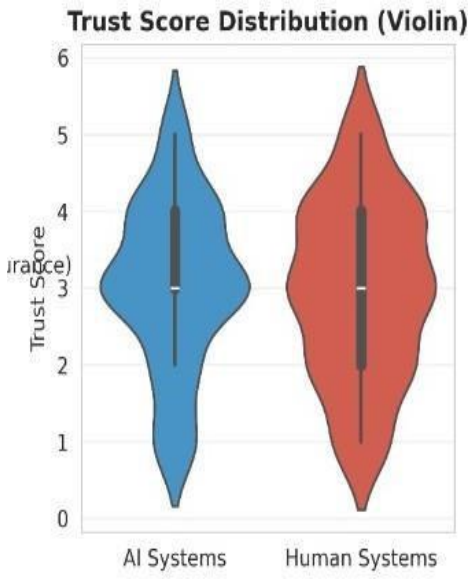
Results and Conclusion:

Results:

The findings demonstrate that AI-based and human-operated fraud detection systems are rated as moderately trusted by the respondents. As both of the two respected systems had identical mean scores and which indicates that participants often tend towards the regard that both of the two methods are equally reliable.

But the tendency of the answers was not exactly similar. Confidence in human-managed systems was more fractured. Approximately, 35 percent of the respondents reported high trust and 34.2 percent of the surveyed said they had low trust. On the contrary, the confidence of AI systems was more focused on the neutral category, 41.9% of respondents chose a neutral answer. Such a distinction implies that human systems form more vigorous opinions, positive, and negative, whereas AI systems are perceived in a more ambivalent or sceptical manner.





In answers to the question of what they would do in the instance of the suspected fraud, 53.45% of the answer was that they would immediately block the account using AI. In the meantime, 46.55% of them wanted to be contacted by a human representative with a verification call. The trust scores were comparable, though; a little more participants with the faster automated response.

Conclusion:

According to the results of the ANOVA there is not much difference in trust levels between different age groups in either system ($p = 0.913$ and $p = 0.584$). It implies that a general trust of the fraud detection methods among the various age groups is relatively similar.

Although there is statistical equality in the trust score levels the results suggest that there is a slow change in the liking of behaviour. Young aged respondents are tending slightly more towards the AI-based immediate blocking systems, which may be explained by its speed and efficiency. The older respondents however favour human-based verification protocols, probably due to reassuring and transparency.

ISSN: P-2455-0515
E- 2394-8450

Educreator Research Journal

OPEN ACCESS

Volume–XIII, Special Issues–I(a) Jan – Feb 2026

Original Research Article

Finally, the research concludes that the AI technology has achieved trust parity with human intervention in the banking fraud detection. Although the two systems are equally trusted, consumer behaviour particularly among the younger users implies that there is increased acceptance of automated solutions.

References:

1. *Ali, A., et al. (2024). The shift to Industry 5.0: Collaborative intelligence and human-machine synergy in modern systems. Journal of Industrial Transformation.*
2. *McNally, R., & Bastos, D. (2025). The trust paradox: Addressing the 'black-box' problem in automated financial decision-making. International Journal of Banking & Ethics.*
3. *Venu, S. (2025). Technical foundations of AI-driven fraud detection: Scalability and pattern recognition in high-volume transaction data. Fintech Research Quarterly.*
4. *Yeo, J., et al. (2025). Human-in-the-Loop (HITL) architectures for ethical oversight in high-risk financial AI systems. Global Review of Financial Technology.*

Cite This Article:

Pabarekar M.V., Singh V.R., Pagade A.D. & Momin A.S. (2026). *Human-in-the-Loop Vs. Lights-out Automation.* **In Educreator Research Journal: Vol. XIII (Issue I),** pp. 69–76.

Doi: https://doi.org/10.5281/zenodo.19881557