WWW.IJRPR.COM

International Journal of Research Publication and Reviews

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

Leveraging Customer Trust in Financial Services through AI Chatbots, A Study with Reference to Private Sector Banks

Ms. A. Bhagya Sri [1], Dr. Narendra [2]

[1]Research Scholar, BBA Department, KLEF, Green Fields, Vaddeswaram, Guntur District, Andhra Pradesh, India, 522502
Email: bhagyasriagniparthi.29@gmail.com
[2] Research Guide, BBA Department, KLEF, Green Fields, Vaddeswaram, Guntur District, Andhra Pradesh, India, 522502

DOI : https://doi.org/10.55248/gengpi.07.1026.34031

ABSTRACT

AI chatbots are becoming a growing feature in banks' customer support, but their usefulness s relies on consumer confidence in the technology to handle financial information. This research explores the level of trust that customers have in AI chatbots of the private sector banks in India. The structured online questionnaire resulted in 75 responses with 35 respondents (46.7%) having interacted with their bank's chatbot. Data were analyzed using chi-square and one-way ANOVA and multiple linear regression in Microsoft Excel. Chatbot use was significantly associated with gender, $\chi^2(1) = 7.83$, $p = .005$, $\varphi = .330$: 65.6% of men but 32.5% of women had used a chatbot. There was no significant difference in mean trust among the usage frequency groups for users, $F(3, 31) = 1.72$, $p = .183$, $\eta^2 = .143$, with the highest trust reported by the users who used on a daily basis ($M = 4.00$). Overall, the perceived accuracy and ease of use, responsiveness and security together accounted for 67.9% of the variance in trust, $F(4, 30) = 15.83$, $p < .001$. Only security ($\beta = .589$) and ease of use ($\beta = .527$) were significant predictors. The results indicate that private banks can establish trust-building measures with customers by demonstrating robust data security, user-friendliness and flexibility of the chatbots, and bridging the gap for women's adoption. Small sample size and primarily young students, so results are indicative and not generalizable.

The topics covered include AI chatbots, customer trust, financial services, perceived security, ease of use and private sector banks.

1. Introduction

Digital transformation has transformed the way that banks connect with customers. Enquiry on balance, fund transfer, complaint handling and answering product enquiries are now among features handled by conversational AI, typically in the form of a chatbot or a virtual assistant. The private banks in India have been amongst the first to introduce such an assistant on their website or mobile application.

In financial services, customers provide sensitive personal and financial information, and if something is incorrect or the information is misused, customers' money is at stake. Trust is thus at the heart of it all. If a customer does not believe in a chatbot, they will bail on it, ask the same question via a more expensive method or go without digital self-service. Banks that grasp the elements that make up trust will be able to use that understanding to boost adoption, reduce service costs, and improve customer relationships.

1.1 Need and Justification

The private sector banks have poured a lot of resources into the development of AI chatbots, aiming to reduce expenses for customer services and provide 24/7 assistance, but only if customers are willing to rely on the capabilities of these tools for actual, rather than hypothetical, financial transactions rather than going to branches or call centres. The current research surrounding the adoption of the chatbots has largely been based on the technology acceptance literature or global markets, neglecting primary research that examines how private sector bank customers actually interact with and trust the technology, specifically in India.

This study finds that less than 50% of the customers surveyed have interacted with their bank's chatbot, there is a definite gender disparity in adoption and, in fact, there are statistical indicators that point to the characteristics of the service that customers find more trustworthy, providing private sector banks with priorities to act on instead of a vague directive to “improve the chatbot.” It also provides a robust model of a chatbot trust specific to India to be developed in future research.

1.2 Background of the Study

Private companies' banks are investing in chatbots but not evidenced by their customers' usage or trust. It remains to be seen which services have the greatest impact on trust, whether there are differences in trust depending on how often customers interact with the chatbot, and whether customer segments are evenly distributed in terms of chatbot use.

1.3 The objectives of the study are outlined below:

•To see if there is any correlation between the use of bank chatbots and gender of the customer.

•To analyse the level of trust between groups of customers who use the chatbot differently.

To measure the impact of perceived accuracy, ease of use, responsiveness and security on customer trust.

•To offer recommendations to private sector banks for building and leveraging customer trust in AI chatbots.

1.4 Scope and Significance

The study is based on the 1000 customers who completed an online survey from private sector banks. The responses from those who have a main bank that is a public sector bank are not included in this scope. It provides primary evidence that bank managers can act on concerning trust drivers, as well as the gap between customers that use chatbots and those that don't.

2. Review of Literature

2.1 Trust

The definition of trust by Mayer, Davis and Schoorman (1995) is based on the trustor's willingness to be vulnerable to the other party due to expectations of the trustee's ability, benevolence and integrity. Morgan and Hunt (1994) locate trust as the focal point of prosperous relationship marketing, correlating it with commitment and cooperation. These are the same principles that he or she must live by as a banker, because customers must become vulnerable when it comes to financial information.

2.2 Trust in Technology and Online Services

Davis (1989) demonstrated that perceived ease of use and usefulness are important elements in the technology acceptance process, and Venkatesh et al. (2003) incorporated other models into the Unified Theory of Acceptance and Use of Technology. These models were extended by Gefen, Karahanna and Straub (2003) and Pavlou (2003), who demonstrated that trust and perceived risk have significant impact on online transaction intentions. Similarly, Parasuraman, Zeithaml and Malhotra (2005) found that there were four key components of electronic service quality: ease of use, responsiveness, accuracy and security, equivalent to efficiency, system availability, fulfilment and privacy respectively.

2.3 Faith in Artificial Intelligence and Chatbots

People have been willing to attribute understanding to machines in early conversational programs like ELIZA (Weizenbaum 1966). Følstad and Brandtzaeg (2017) explain how chatbots impacted on human-computer interaction and made services conversational. Lee and See (2004) suggest that the use of automation must be matched with appropriate levels of trust, and Glikson and Woolley (2020) provide an overview of empirical research that indicates that trust in AI is influenced by tangibility, transparency, reliability and immediacy behaviours of the system. These studies collectively indicate that the trust in financial chatbots is likely to be influenced by perceived performance, usability, and security.

2.4 Research Gap

The majority of the research on AI trust happens in the West and is broadly applicable. There is limited evidence on customer trust in private sector banks' chatbots, and who is using them. This study fills that void. Include 5-10 studies that your guide has suggested, and specifically related to Indian and banking economics.

3.The conceptual framework and hypotheses.

The dependent variable is customer trust in the bank chatbot. Perceived accuracy, ease of use, responsiveness and security are the independent variables. Some of the customer features that can be related to chatbot use or trust include gender and usage frequency. The following hypotheses are tested:

$H1_0$: There is no significant association between gender and use of the bank chatbot.

$H2_0$: No significant difference between the means of the trust on the usage frequency of chatbots.

$H3_0$: Trust is significantly and positively influenced by the perceived accuracy (H3a), ease of use (H3b), responsiveness (H3c) and security (H3d).

4. Research Methodology

4.1 Research Design

A descriptive and explanatory research design and cross-sectional survey was used.

4.2 Collecting samples and data

The type of data collected was primary data that was gathered using an online structured questionnaire (Google Forms) using convenience sampling. 117 responses were received in total (Distributed via form). The study is based on the private sector banks, so 40 respondents with main bank as public sector bank and 2 respondents who did not have a bank were excluded and 75 respondents were selected for this study. The primary bankers of these customers are HDFC Bank (49.3%), Axis Bank (28.0%), Kotak Mahindra Bank (9.3%), ICICI Bank (6.7%) and four other private banks (6.7%). Questions of trust were only included in the survey for the users who responded yes (n = 35, 46.7%), with the non-users (n = 40, 53.3%) being the analytical sample for ANOVA and regression. The chi-square test for bots in the private sector is provided for those who identified as male or female.

4.3 Instrument and Measures

The questionnaire contained an attitude section with 5 point likert scale from 1 (very disagree) to 5 (very agree), and a profile section with the following questions: age of the respondents, gender, education, occupation, bank used by the respondent, use of chattbot, how often the respondent used the chatbot, what services the chatbot provides to the respondent. The items for perceived accuracy (Q9), ease of use (Q10), responsiveness (Q11) and security (Q12) were each measured once. The overall measure of trust was the average of four items: trust in the balance enquiry (Q14); trust in fund transfers (Q15); trust in complaint handling (Q16); and overall trust (Q18). The Cronbach's alpha coefficient for the trust scale was .780, which is satisfactory for this scale (Nunnally, 1978). For descriptive purposes, trust was broken down into three categories: Low (<3.0); Medium (3.0 to below 4.0); and High (4.0 and above). The comfort in sharing account details (Q13), preference for a human employee (Q17) and likelihood of recommending the chatbot (Q19) are reported descriptively.

4.4 Statistical Tools

Data were analyzed using MS Excel. Data were summarised using descriptive statistics and Pearson correlations. A two-way chi-squared test for independence was used to compare the factors of gender and use of chatbots. A one-way ANOVA was used to compare the mean trust between the various usage frequency groups with Bonferroni adjusted pairwise comparisons. Multiple linear regression was used to estimate the effect of the four predictors on the outcome of trust. The level of significance was 0.05. The age was not used as a test variable because the age group 18-25 was large enough (80.0%) to be valid for testing.

5. Data Analysis and Interpretation

5.1 Respondent Profile

Table 1: Profile of respondents from the private sector banks (n = 75)

Variable	Category	Count	Percent
Age group	Below 18	9	12.0%
	18-25	60	80.0%
	26-35	2	2.7%
	36-50	2	2.7%
	Above 50	2	2.7%

Variable	Category	Count	Percent
Gender	Male	32	42.7%
	Female	40	53.3%
	Prefer not to say	3	4.0%
Education	Secondary / Higher Secondary	22	29.3%
	Undergraduate	3	4.0%
	Graduate	41	54.7%
	Postgraduate	6	8.0%
	Professional	3	4.0%
Occupation	Student	64	85.3%
	Salaried employee	5	6.7%
	Business / self-employed	5	6.7%
	Homemaker / Retired	1	1.3%
Main bank	HDFC Bank	37	49.3%
	Axis Bank	21	28.0%
	Kotak Mahindra Bank	7	9.3%
	ICICI Bank	5	6.7%
	IDFC First Bank	2	2.7%
	Bandhan, City Union, Federal (1 each)	3	4.0%
Used bank chatbot	Yes	35	46.7%
	No	40	53.3%

Source: Survey data.

The sample is predominantly young and predominantly made up of students (80.0% and 85.3% respectively). Women make up 53.3% of the respondents. The sample consists of customers from both HDFC Bank and Axis Bank, accounting for 77.3% of the total. Less than half (46.7%) had interacted with a chatbot from their bank.

Table 2 shows a profile of the respondents who used a chatbot (n = 35).

Variable	Category	Count	Percent
Usage frequency	Daily	9	25.7%
	Weekly	8	22.9%
	Monthly	8	22.9%
	Rarely / once	10	28.6%

Variable	Category	Count	Percent
Services used*	Balance enquiry / statements	16	45.7%
	Fund transfer	16	45.7%
	Complaint handling	13	37.1%
	Loan or card queries	9	25.7%
Trust level	Low (< 3.0)	5	14.3%
	Medium (3.0 to < 4.0)	15	42.9%
	High (≥ 4.0)	15	42.9%

Source: Survey data

The frequency distribution is fairly even with little or none most frequently and once the largest (28.6%). The top uses were balance enquiries/consultation and fund transfer (45.7% each). The level of trust is moderate, with 42.9% in the High band, 42.9% in the Medium band and 14.3% in the Low band.

5.2 Descriptive Statistics and Correlations

Table 3 presents item and construct statistics for the users (n = 35).

Statement	Mean	Std. Dev.	% Agree
Q1 Accurate answers (Accuracy)	3.26	0.85	37.1%
Q2 Easy to use (Ease of use)	3.74	0.78	60.0%
Q3 Responds quickly (Responsiveness)	3.89	0.80	68.6%
Q4 Data is safe (Security)	3.63	0.77	62.9%
Q13 Comfortable sharing account details	3.43	1.12	51.4%
Q17 Prefer a human for important matters	3.91	0.92	74.3%
Q19 Would recommend the chatbot	3.54	1.01	54.3%
Trust Score (Q14, Q15, Q16, Q18)	**3.60**	**0.77**	**65.7%***

Data from the survey. % Agree = percentage that answered Agree or Strongly agree. Share with a score of 3.5 or higher on the Trust Score.

The highest rating was given to responsiveness (M = 3.89) and the lowest rating was given to accuracy (M = 3.26), with only 37.1% agreeing that the chatbot provides accurate answers. Mean trust was moderately positive (M = 3.60, SD = 0.77). Interestingly, 74.3% of respondents said they would rather deal with a human employee if it was a big issue, and that trust in the chatbot does not compete with the strong preference for human service.

Table 4: Pearson correlation matrix (users)

	Accuracy	Ease	Respons.	Security	Trust	Recomm.
Accuracy	1.000	0.456	0.088	0.374	0.461	0.243
Ease of use	0.456	1.000	0.283	0.130	0.591	0.332

Responsiveness	0.088	0.283	1.000	0.312	0.281	0.153
Security	0.374	0.130	0.312	1.000	0.643	0.494
Trust	0.461	0.591	0.281	0.643	1.000	0.624
Recommend	0.243	0.332	0.153	0.494	0.624	1.000

Source: Survey data.

All correlations are positive. The most significant correlation with trust is found with security (r = .643), followed by ease of use (r = .591), and accuracy (r = .461). There is a strong relationship between trust and willingness to recommend (r = .624). The intercorrelations between the predictors range from .09 to .46 and variance inflation factors are not greater than 1.5, indicating that multicollinearity is not a problem.

5.3 Chi-Square Test: Gender and Use of Bank Chatbot

$H1_0$: There is no significant association between gender and use of the bank chatbot.

Table 5: Gender by Chatbot use – observed counts

Gender	Used chatbot	Not used	Total	% who used
Male	21	11	32	65.6%
Female	13	27	40	32.5%
Total	**34**	**38**	**72**	**47.2%**

Source: Survey data. Three users who do not want to specify gender are not included.

Table 6: The results of the chi-square test

Statistic	Value	Remark
Chi-square (χ^2)	7.827	Calculated
Degrees of freedom	1	(2-1)(2-1)
Critical value ($\alpha = 0.05$)	3.841	Table value
p-value	0.005	< 0.05
Phi (φ)	0.330	Medium effect
Minimum expected count	15.11	Assumption met (≥ 5)

Data collected from surveys; calculations made in EXCEL.

The calculated χ^2 (7.827) exceeds the critical value (3.841) and p = .005 is below .05, so $H1_0$ is rejected. There is a significant association with Chatbot use across gender, with 65.6% of males compared with 32.5% of females using their bank's Chatbot in this sample. The effect is medium as per Cohen's (1988) benchmarks. This result remains after applying a continuity correction (Yates corrected p = .010). The survey didn't gather information on why users didn't use a chatbot, so it is not possible to determine what reasons might explain the difference from the data.

5.4 One-Way ANOVA: Trust by Usage Frequency

$H2_0$: There are no significant differences between the mean trust of the groups representing different frequencies of use of the chatbot.

Table 7: reports the mean trust scores for each usage frequency across the three groups.

Usage frequency	n	Mean trust	Std. Dev.
Daily	9	4.000	0.875
Weekly	8	3.719	0.364
Monthly	8	3.469	0.737
Rarely / once	10	3.250	0.858
Total	**35**	**3.600**	**0.775**

Source: Survey data.

Table 8: ANOVA summary

Source	SS	df	MS	F	p-value
Between groups	2.916	3	0.972	1.723	0.183
Within groups	17.484	31	0.564		
Total	20.400	34			

Source: Survey data. $\eta^2 = 0.143$; F critical (3, 31) at 0.05 = 2.911.

The calculated F = 1.723 is less than the critical F = 2.911 and the p value = .183, therefore $H2_0$ is not rejected, and it means that the mean trust is not significantly different between the usage frequency groups. The effect size ($\eta^2 = .143$) is large, as defined by Cohen (1988) and Mean trust falls steadily from daily users (4.00) to rarely or once users (3.25). The test, however, had low power given that it had only 8–10 users per group, and none of the differences between pairs of user groups reached significance (smallest p = .225 when comparing daily with rarely or once). The variances within groups were not identical (most uniform were users for a week), but a separate Levene test revealed the variances were not significantly different (p = .509), and the assumption of ANOVA is acceptable. More samples is required to determine if increased usage is associated with increased trust.

5.5 Multiple Linear Regression: Drivers of Trust

The factors of perceived accuracy, ease of use, responsiveness and security positively significantly affect trust.

Table 9: The regression model summary

Statistic	Value	Remark
Multiple R	0.824	Strong
R^2	0.679	67.9% variance explained
Adjusted R^2	0.636	
Std. error of estimate	0.468	
F (4, 30)	15.830	$p < .001$

Data from survey; calculated in Excel (LINEST). n = 35.

Table 10: presents the regression coefficients for the dependent variable (Trust Score)

Predictor	B	Std. error	t	p	Std. beta (β)
(Constant)	-0.327	0.558	-0.586	0.562	-
Accuracy	0.004	0.115	0.038	0.970	0.005
Ease of use	0.524	0.121	4.315	< .001	0.527
Responsiveness	-0.051	0.111	-0.457	0.651	-0.052
Security	0.593	0.119	4.980	< .001	0.589

Source: Survey data. Fitted model: Trust = -0.327 + 0.004(Accuracy) + 0.524(Ease) - 0.051(Responsiveness) + 0.593(Security).

The model is statistically significant ($F = 15.83$, $p < .001$) and has an adjusted R^2 of .636, accounting for 67.9% of the variance in the criterion, trust. Only two predictors are significant: When the rest of the factors are held constant, a one-point increase in perceived security contributes 0.593 points to trust and a one-point increase in ease of use contributes 0.524 points to trust. The standardised effects for both of them are comparable (β = .589, .527). The effect of the two is not independent and significant. Thus, H3b and H3d are supported, while H3a and H3c are not supported. The number of observations (n=35) is at or above Hair et al's (2019) suggested minimum ratio of 5 observations per predictor, but not at the preferred ratio, so the coefficients should be used with caution.

5.6 Open-Ended Responses

There were 35 users who answered the open question on the biggest problem with bank chatbots, with 23 of those users answering the question. 15 responded that there was no problem. Seven boundaries of the Chatbot understanding are defined: it does not understand the actual problem, provides only defined answers, repeats similar answers, provides wrong or irrelevant answers. According to one of the respondents the chatbot is not to be trusted. The lack of understanding and rigid menus were also mentioned as problems by customers and is reflected in the relatively low accuracy rating (M = 3.26).

6. Summary of Finding

Hypothesis	Test	Result
$H1_0$	Chi-square: gender and chatbot use	Rejected (p = .005)
$H2_0$	One-way ANOVA: trust by usage frequency	Not rejected (p = .183)
H3a	Regression: accuracy → trust	Not supported (p = .970)
H3b	Regression: ease of use → trust	Supported (p < .001)
H3c	Regression: responsiveness → trust	Not supported (p = .651)
H3d	Regression: security → trust	Supported (p < .001)

Less than half of private sector banks (46.7%) had used their bank's chatbot and men were more likely to do so than women (65.6% to 32.5%).

Trust of users is moderately positive (mean 3.60 on 5); the percentages of high and low trust are 42.9% and 14.3% respectively.

Security and ease of use are the major factors that influence trust, whereas accuracy and responsiveness do not influence trust once they are considered.

The mean trust does not vary significantly according to frequency of use, with highest scores being given by those who use it daily.

Still, almost three quarters (74.3%) prefer dealing with a human employee for sensitive situations.

7. Discussion

The impact of perceived security is in line with e-service quality research which considers privacy and security as fundamental components (Parasuraman et al., 2005) and with the findings that perceived risk has an impact on online trust (Pavlou, 2003). This is a significant effect of ease of use, which is consistent with TAR (Davis, 1989; Venkatesh et al., 2003). Security and ease of use are not only equally weighted, but both are important.

Trust alone is a significant predictor of accuracy (r = .461), but it is not significant in the regression. This may be due to the fact that it is moderately intercorrelated with accuracy (r = .456) and security (r = .374) and does not contribute much if these are included. This is not to say that accuracy is not important: it was the lowest rated attribute (M = 3.26) and the open-ended complaints were primarily of wrong, irrelevant or repetitive answers. According to Lee and See (2004), the key is to build trust to reflect the actual performance, therefore ongoing accuracy is still significant for trust to be sustained. The highest rating (M = 3.89) was given for the responsiveness, but this did not predict trust, which could suggest that responsiveness is what is expected of a chatbot, but not necessarily what leads to trust.

The gender gap in the use of the chatbot is significant, but the survey does not shed light on it. It could be due to awareness differences, or banking practices, or how the sample was reached, and needs further research. The failure to find any significant usage effect on trust should be interpreted with caution as it occurred with small groups.

8. Managerial implications and recommendations

Security is the strongest driver of trust: be visible with clear data-protection messages, with strong authentication for sensitive actions and with explanations of how data is used.

Make it easier and more flexible: enable free text answers, avoid too many menu options and provide standard answers when respondents were identified as issues.

Provide a fast transfer to the human agent on complex or high-value issues as most users still like to have a human agent helping them with important matters.

Targeted communication and guided first use to raise awareness and confidence building among female customers to broaden the base of customers who can trust the chatbot.

Use actual queries for training to improve the accuracy of answers (accuracy was the lowest rated attribute).

9. Limitations and scope for further research.

The findings are not generalizable as the study is a small convenience sample and just 35 private sector bank customers had interacted with a chatbot. The sample was composed of a high proportion of young students (80.0% aged 18-25 and 85.3% students) and 9 were under 18 years, which meant that this prevented the assessment of age and occupation effects, although some of these were tested in the results. The results are largely based on customers of HDFC Bank and Axis Bank as they make up 77.3% of the sample. The results do not apply to customers of public sector banks. There are both self-reported and cross-sectional responses, which restrict causal inferences. Single items were used for measuring the predictors and the survey failed to investigate the reasons for avoiding using chatbots by non-users. Future research can be conducted with larger and wider sample, multi-item scales, structural equation modelling, comparing the differences between private and public sector banks, and longitudinal designs to monitor the development of trust over time.

10. Conclusion

The findings of this study reveal that the trust in AI chatbots among the customers of the private sector banks is at moderate level and is influenced primarily by perceived security and ease of use of the AI chatbots. The use of chatbots is not evenly distributed – women are significantly less inclined to use a chatbot than men – and many customers still wish to use human service for major issues. To build trust with customers by using chatbots, private sector banks need to enhance the presence of security measures, improve the ease and flexibility with which they can interact with the bot and ensure that there are clear pathways to human support.

References

Cohen, J. (1988). This is the second edition of Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates.

Davis, F. D. (1989). The acceptance of information technology, perceived usefulness and perceived ease of use. MIS Quarterly, 13(3), 319-340.

Følstad, A. & Brandtzaeg, P. B. (2017). Chatbots and the new world of HCI. Interactions, 24(4), 38-42.

Gefen, D., Karahanna, E., & Straub, D. W. (2003). A combined model of trust and TAM in online shopping. MIS Quarterly, 27(1), 51-90.

Glikson, E., & Woolley, A. W. (2020). Empirical studies: Human trust in artificial intelligence. Academy of Management Annals, 14(2), 627-660.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate data analysis (8th ed.). Cengage Learning.

Lee, J. D., & See, K. A. (2004). Reliance on automation – Designing for. Human Factors, 46(1), 50-80.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative approach to organizational trust. Academy of Management Review, 20(3), 709-734.

Morgan, R. M., & Hunt, S. D. (1994). The relationship marketing commitment-trust theory. Journal of Marketing, 58(3), 20-38.

Nunnally, J. C. (1978). The second edition of the Psychometric theory. McGraw-Hill.

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). A multiple-item instrument that measures electronic service quality (E-S-QUAL). Journal of Service Research, 7(3), 213-233.

Pavlou, P. A. (2003). Integrating trust and risk with the technology acceptance model in the context of consumer acceptance of electronic commerce. Educational Electronic Commerce, 7(3): 101-134.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). Getting to a single perspective on user acceptance of information technology. MIS Quarterly, 27(3), 425-478.

Weizenbaum, J. (1966). A computer program designed for the study of natural language communication between man and machine. Communications of the ACM, 9(1), 36-45.