

AI vs. Humans in U.S. Retail Banking: A Pilot Study on Customer Satisfaction and Service Excellence

Sifat Mahmud 

Purpose. This pilot study aims to explore the perceived effectiveness of artificial intelligence (AI) and human interaction models within a university-affiliated sample, focusing on U.S. retail banking customer satisfaction and service excellence. **Design / Method / Approach.** The pilot study involved 50 participants from a university community—U.S. retail banking customers affiliated with Wright State University. A structured survey was conducted, and data were analyzed using descriptive statistics, predictive analytics, ANOVA, regression analysis, and t-tests to assess satisfaction levels, response times, and service preferences. **Findings.** Results indicate that human interaction consistently outperformed AI, with an average satisfaction score of 4.14 compared to 3.56 for AI ($p < 0.05$). Male participants rated AI higher, whereas females preferred human interaction. Regression analysis revealed that AI satisfaction was primarily influenced by service consistency ($p = 0.035$, $R^2 = 0.176$), while human satisfaction was driven by personalized service ($p = 0.009$, $R^2 = 0.220$). These results suggest that empathy and personalization remain central to service excellence, while AI's consistency can enhance operational efficiency. **Theoretical Implications.** The findings contribute to understanding the human–AI service trade-off by integrating behavioral and demographic dimensions into service design. **Practical Implications.** The study recommends enhancing the consistency of AI-driven systems and investing in employee training programs to strengthen empathy and personalization, which together foster loyalty and customer trust. **Originality / Value.** This work empirically substantiates the dual path toward technological efficiency and human-centric empathy, highlighting hybrid AI–human approaches as optimal for banking economics. **Research Limitations / Future Research.** Limited sample size and homogeneity restrict generalization; future research should employ larger, demographically diverse samples and longitudinal designs to explore mediating factors such as trust and cultural context. **Article Type.** Exploratory Research / Pilot Study.

Keywords:

artificial intelligence, human interaction, customer satisfaction, retail banking, service excellence, hybrid AI-human models

Мета. Це пілотне дослідження спрямоване на вивчення сприйнятої ефективності моделей взаємодії штучного інтелекту (ШІ) та людини у вибірці, пов'язаній з університетським середовищем, із фокусом на задоволеності клієнтів роздрібного банкінгу США та досконалості обслуговування. **Дизайн / Метод / Підхід.** Пілотне дослідження охопило 50 учасників університетської спільноти — клієнтів роздрібного банкінгу США, пов'язаних з Університетом Райта. Проведено структуроване анкетування, а отримані дані проаналізовано із застосуванням описової статистики, предиктивної аналітики, ANOVA, регресійного аналізу та t-тестів для оцінювання рівнів задоволеності, часу реагування та уподобань у сервісних моделях. **Результати.** Виявлено, що взаємодія з людьми стабільно перевищує ефективність ШІ: середній бал задоволеності становив 4,14 проти 3,56 для ШІ ($p < 0,05$). Чоловіки оцінювали ШІ вищими балами, тоді як жінки надавали перевагу людській взаємодії. Регресійний аналіз показав, що задоволеність ШІ визначається насамперед послідовністю обслуговування ($p = 0.035$, $R^2 = 0.176$), тоді як людська задоволеність — персоналізованим підходом ($p = 0.009$, $R^2 = 0.220$). Отримані результати свідчать, що емпатія та індивідуалізація залишаються ключовими для високої якості сервісу, а послідовність ШІ сприяє підвищенню ефективності процесів. **Теоретичне значення.** Дослідження поглиблює розуміння компромісу між людським і ШІ-обслуговуванням, враховуючи поведінкові та демографічні чинники у проектуванні сервісів. **Практичне значення.** Рекомендовано підвищити послідовність роботи систем ШІ та інвестувати у підготовку персоналу для розвитку емпатії й персоналізованого обслуговування, що разом зміцнює довіру та лояльність клієнтів. **Оригінальність / Цінність.** Робота надає емпіричні підтвердження доцільності поєднання технологічної ефективності з людською емпатією, підкреслюючи переваги гібридного підходу «людина – ШІ» для банківського сектору. **Обмеження дослідження / Майбутні дослідження.** Обмеженість вибірки та її однорідність знижують рівень узагальнення результатів; подальші дослідження мають охопити ширші та різноманітніші вибірки із застосуванням лонгitudinalних методів і вивченням таких чинників, як довіра та культурний контекст. **Тип статті.** Прикладне дослідження / Пілотне дослідження.

Ключові слова:

штучний інтелект, людська взаємодія, задоволеність клієнтів, роздрібний банкінг, досконалість обслуговування, гібридні моделі ШІ–людина

Contributor Details:

Sifat Mahmud, M.S. in Marketing Analytics and Insights, Master of Professional Banking, Wright State University: Dayton, Ohio, US, mahmud.21@wright.edu

The exaggerated adoption of Artificial Intelligence (AI)–driven technologies in U.S. retail banking represents a double-edged sword for the national economy (McKinsey & Company, 2021). In fact, businesses rely on AI with the expectation that it will boost efficiency and significantly reduce operational costs; however, a key concern remains: the ultimate impact on customer satisfaction. This question is particularly relevant in areas where human touch, empathy, personalization, and understanding are essential (Deloitte Network, 2024). This research investigates the effectiveness of AI versus human interaction in retail banking, focusing on their influence on customer satisfaction and service excellence. Additionally, this study addresses a persistent concern for the U.S. economy: harmonizing technological advancement with customer-centric service delivery.

Literature Review

The increasing adoption of AI technologies in retail banking has fundamentally transformed service delivery models (Scheffler & Puczyk, 2025). While AI implementations demonstrate measurable improvements in operational efficiency and cost reduction (Patnayak, 2021), customers continue to express a preference for human interaction in complex service scenarios (Huang & Rust, 2022). This preference is particularly evident in emotionally sensitive banking interactions, such as loan applications and financial counseling (Huang & Rust, 2022). Current research identifies key demographic variations in technology acceptance, with younger male customers showing greater comfort with AI-driven services compared to older age groups (Méndez-Suárez et al., 2023). Gender differences are especially pronounced, with female customers consistently rating human interactions higher across all age ranges (Ameen et al., 2021). These findings align with broader service literature emphasizing the importance of emotional intelligence in customer satisfaction (The Contentstack Team, 2024).

AI systems currently excel in routine transactions but face challenges in addressing nuanced customer needs that require empathy and judgment (Huang & Rust, 2022). Emerging hybrid models that strategically combine AI efficiency with human empathy show promise for optimizing service delivery and customer satisfaction (Huang & Rust, 2022).

Methodology

Data Collection

A survey was conducted among participants from Wright State University's Raj Soin College of Business, focusing on key metrics such as customer satisfaction, response time for AI and human interaction, and service preferences. The survey combined quantitative measures with qualitative feedback to capture personal experiences where AI-driven technology and human interaction play roles in the U.S. retail banking system.

Data Cleaning

Survey responses were rigorously cleaned to eliminate incomplete or irrelevant entries, ensuring the reliability of the results. Numerical responses were coded to facilitate statistical analysis. Insights derived from the analysis informed actionable recommendations aimed at supporting economic efficiency in the future of the U.S. retail banking industry.

Analytical Approach

To examine the effectiveness of AI versus human interaction in U.S. retail banking, a comprehensive five-stage analytical approach was employed. This approach integrates descriptive and inferential statistics with data visualization techniques to systematically evaluate customer satisfaction, response times, service preferences, and their broader economic implications.

Descriptive Statistics were first applied to summarize central tendencies (means, standard deviations) and distributions of key variables, including customer satisfaction trends, response times, and preference distributions. This stage provided foundational insights into service dynamics and allowed identification of patterns in the data without making inferential claims, forming the basis for subsequent analyses.

ANOVA (Analysis of Variance) was used to compare mean satisfaction levels between AI and human interactions. The F -statistic, calculated as

$$F = \frac{\text{between-group variance}}{\text{within-group variance}},$$

was employed under the null hypothesis of equal means, with assumptions of normality and homogeneity of variance verified using Levene's test. This analysis identified the optimal balance of service delivery to enhance customer loyalty and informed recommendations on critical factors contributing to economic stability in the banking sector.

Regression Analysis examined the relationship between response time and customer satisfaction. Multiple linear regression models were fitted, e.g.,

$$\text{Satisfaction} = \beta_0 + \beta_1(\text{Response Time}) + \varepsilon,$$

where β coefficients represent the change in satisfaction per unit change in the predictor. Assumptions of linearity, independence, homoscedasticity, and normality of residuals were verified using scatterplots, Durbin-Watson tests, and Q - Q plots. The coefficient of determination R^2 quantified the variance explained, while p -values assessed the significance of each predictor at $\alpha = 0.05$. This stage provided insights into how operational competence impacts economic outcomes in the banking sector.

t -tests were conducted to compare preferences for empathy and personalization between AI and human services. Independent samples t -tests, calculated as

$$t = \frac{M_1 - M_2}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}}$$

assuming equal variances confirmed by an F -test, were complemented with effect size calculations (Cohen's d) to evaluate practical significance. This analysis highlighted the economic importance of trust and relationship-building in financial markets and quantified differences in customer perception between AI and human interactions.

Finally, **Data Visualization** consolidated graphical and tabular representations of all findings, including bar charts (e.g., Figure 1 for satisfaction by demographics), which facilitated clear interpretation of trends and supported data-driven decisions to improve service delivery and economic efficiency. All analyses were conducted in R (version 4.3.1), with significance defined at $p < 0.05$. The survey instrument is detailed below.

Self-Reflection on Research Scope

The present pilot study was designed to explicitly lay the groundwork for foundational insights and a methodological proof-of-concept in a specific university context. While future research would benefit from an enlarged and more diverse sample, including additional variables such as 'trust,' and more extensive statistical analyses, these elements were beyond the practical and conceptual scope of this initial investigation. The present work thus represents a critical first step, laying necessary groundwork for more extended research programs.

Survey Questionnaire

The following questionnaire aims to capture participants' experiences and perceptions of AI and human interactions in U.S. retail banking.

1. Demographic Information

- Age group: 18–22 / 22–29 / 30–39.
- Gender: Male / Female.

2. Please rate the following statements on a scale of 1–5 (1 = Strongly Disagree, 5 = Strongly Agree):

- Q1_a: AI is efficient in handling banking queries.
- Q1_b: Human representatives provide personalized service.
- Q1_c: AI responses are consistent.
- Q1_d: Human representatives demonstrate empathy.
- Q2: I am satisfied with AI interactions overall.
- Q3: I am satisfied with human interactions overall.
- Q4_a: Human interactions offer convenience and speed.
- Q4_b: Human interactions provide personalized service.
- Q4_d: Human interactions demonstrate empathy and

understanding.

- Q5_a: AI response time is satisfactory.

3. Open-ended question:

- Describe your experiences with AI and human banking interactions.

Results

This section presents the findings of the study, organized according to the three research questions. The analysis examines customer satisfaction with AI and human services, evaluates feedback on service efficiency and responsiveness, and explores preferences regarding empathy, personalization, and human-centric engagement in U.S. retail banking. The following subsections summarize both quantitative and qualitative insights from the survey, highlighting patterns in satisfaction, performance, and preferences.

Research Question 1: Customer Satisfaction

This section evaluates overall customer satisfaction with both AI-driven and human interactions in U.S. retail banking services.

Based on survey responses, a comparative analysis was conducted to examine satisfaction levels with AI-enabled automated services and human-assisted interactions across different age and gender groups.

The results indicate a consistent pattern: human interactions were generally rated higher than AI interactions, regardless of demographic category. As shown in Figure 1, the bar chart presents the mean satisfaction scores (on a 1–5 scale) for AI (Q2) and human (Q3) interactions across age–gender subgroups, demonstrating the persistent advantage of human engagement in perceived service quality.

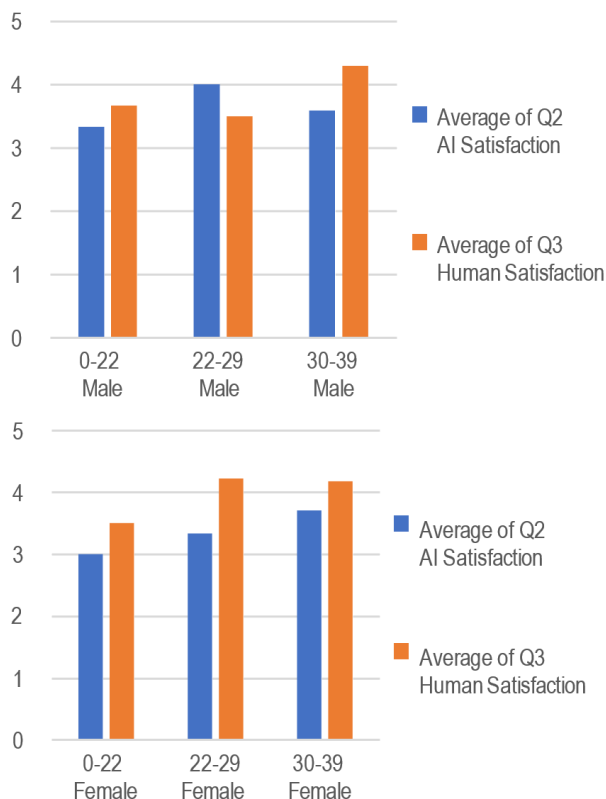


Figure 1 – Measuring Customer Satisfaction: Evaluate overall satisfaction with both AI and human interaction (Source: Author)

The analysis of survey data reveals clear demographic patterns in perceived satisfaction with AI-based and human-based services in U.S. retail banking. The results are summarized across three analytical dimensions: **age-based movements**, **gender-based variances**, and a **grand total summary** supported by regression analysis.

1. Age-Based Movements

For respondents aged 0–22, the overall satisfaction score for AI services was 3.2, slightly lower than the 3.6 recorded for human interaction. Within this group, males rated both AI (3.33) and human

(3.6) interactions higher than females (3.0 for AI and 3.5 for human).

Among participants aged 22–29, a more distinct satisfaction gap emerged. The average AI satisfaction score was 3.45, while human interaction reached 4.09. Males in this group expressed greater satisfaction with AI (4.0) than females (3.33), whereas females reported substantially higher satisfaction with human interaction (4.22) compared to males (3.5).

For the 30–39 age group, AI satisfaction increased slightly to 3.65, but human satisfaction remained higher at 4.24. Here, males rated human interaction at 4.29 and females at 4.18, while their evaluations of AI services were nearly identical (3.59 and 3.71 respectively).

2. Gender-Based Variances

Across all age categories, males generally rated AI-driven services higher than females, with one exception: in the youngest group (0–22), females expressed slightly higher satisfaction with human interaction. Conversely, females consistently rated human interaction higher in both the 0–22 and 30–39 cohorts, confirming a gender-linked tendency toward valuing empathy and personalization.

3. Grand Total Summary

When aggregated across all demographic segments, the mean satisfaction score for AI services was 3.56, compared with 4.14 for human interactions. This overall difference demonstrates a statistically and practically meaningful preference for human-based customer service in retail banking environments.

The regression model presented in Table 1 (Regression Statistics for AI Satisfaction) further supports these findings. The model yielded $R^2 = 0.176$, indicating that 17.6% of the variance in AI satisfaction can be explained by the predictors included. The analysis was statistically significant ($F(3,46) = 3.28, p = 0.029$), with consistency emerging as the most influential predictor ($\beta = 0.176, p = 0.035$). This suggests that the perceived reliability of AI systems plays a key role in shaping overall satisfaction levels.

The statistical outcomes reveal a consistent trend indicating that human interactions are generally evaluated more positively than AI-based interactions, with an average difference of 0.58 points across all respondent groups. Notable variations in satisfaction emerge when disaggregated by age and gender, particularly within the 22–29 age group, where males tend to assign higher satisfaction scores to AI interactions, whereas females provide higher ratings for human interactions. Overall, males demonstrate slightly greater satisfaction with AI services, while females describe more favorable experiences with human-assisted interactions.

Research Question 2: Customer Feedback Assessing

To assess customer feedback, the study examined how AI and human interactions are perceived in terms of satisfaction, efficiency, consistency, and response time, providing a basis for evaluating overall experience.

Based on the survey responses, regression analysis was performed on AI satisfaction, efficiency, consistency, and response time. This analysis provided a better image on overall experience. Consistency significantly improves satisfaction, while efficiency shows a marginal impact, and response time has no meaningful effect. The AI model ($R^2 = 17.6\%$) was statistically significant overall ($p = 0.029$). R^2 (17.6%) clarifies that 18% of the differences in customer satisfaction with AI are explained by this model. The significance F (0.029) indicates that the overall model is statistically significant, meaning the three factors—Efficiency, Consistency, and Response Time influence—combined have a real impact on satisfaction. **Key takeaways** are that consistency is the most important factor for improving customer satisfaction, efficiency might also help but its impact is less clear because it is not statistically significant in this analysis, and response time does not seem to affect satisfaction much, based on this data.

The analysis of the impact of Personalization and Empathy on satisfaction with human interaction identified empathy as marginally significant, while personalization showed no meaningful impact.

The analysis examines how human personalization and human empathy impact satisfaction with human interaction. The regression model explains 8.1 % of the variation in satisfaction ($R^2 = 0.081$), but the overall model is not statistically significant ($F = 0.137, p >$

0.05), indicating that the combination of Personalization and Empathy does not reliably predict satisfaction at a conventional significance level. As shown in Table 2 (Regression Statistics for Human Satisfaction), empathy demonstrates marginal significance ($\beta = 0.248, p = 0.054$), suggesting a slight positive effect on satisfaction, whereas the negative coefficient for personalization indicates a

potential adverse effect and warrants further investigation into multicollinearity or contextual factors.

Key Takeaways. Empathy might have a slight positive effect on satisfaction, while personalization seems to have a negative effect on satisfaction.

Table 1 – Regression Statistics for AI Satisfaction (Source: Author)

Regression Statistics	
Multiple R	0.419580592
R Square	0.176047874
Adjusted R Square	0.122311865
Standard Error	0.632294157
Observations	50

ANOVA

	df	SS	MS	F	Significance F
Regression	3	3.929388538	1.309796179	3.276162099	0.029264788
Residual	46	18.39061146	0.399795901		
Total	49	22.32			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	1.771441719	0.675844419	2.621079154	0.011840515	0.411037462	3.131845976	0.411037462	3.131845976
Q1 a - AI Efficiency	0.18267135	0.094086339	1.941528937	0.058333969	-0.006714628	0.372057328	-0.006714628	0.372057328
Q1 c - AI Consistency	0.176184345	0.08115173	2.171048524	0.035123112	0.012834384	0.339534306	0.012834384	0.339534306
Q5 a - AI Response Time	0.156891371	0.140985822	1.112816659	0.271568881	-0.126898369	0.44068111	-0.126898369	0.44068111

Table 2 – Regression Statistics for Human Satisfaction (Source: Author)

Regression Statistics	
Multiple R	0.284613343
R Square	0.081004755
Adjusted R Square	0.041898575
Standard Error	0.713282139
Observations	50

ANOVA

	df	SS	MS	F	Significance F
Regression	2	2.107743733	1.053871866	2.07140544	0.137360451
Residual	47	23.91225627	0.50877141		
Total	49	26.02			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	3.751624884	0.539195357	6.957821203	9.47073E-09	2.66690374	4.836346027	2.66690374	4.836346027
Q1 b - Human Personalization	-0.148096565	0.118110212	-1.253884503	0.216084969	-0.385703663	0.089510534	-0.385703663	0.089510534
Q1 d - Human Empathy	0.247910864	0.125371175	1.977415165	0.053876103	-0.004303409	0.500125136	-0.004303409	0.500125136

Comparative Analysis of AI and Human Interaction – Customer Satisfaction includes metrics comparison, predictor analysis, and ANOVA tables. Table 3, Figure 2 summarizes model fit, highlighting AI's superior predictive power (higher R^2 and significance). The R -squared values indicate that AI interactions clarify 17.6% of the variance in satisfaction levels, while human interactions clarify only 8.1%. This indicates that the AI model is better at predicting satisfaction. The adjusted R -squared values show that AI still performs better than the human model. The significance F -values indicate that the AI model has a statistically significant overall model ($p < 0.05$), whereas the human models are not statistically significant ($p = 0.1374$).

Table 3 – Summary of Model Fit Metrics (Source: Author)

Metric	AI Value	Human Value
R-Squared	0.176	0.081
Adjusted R-Squared	0.1223	0.0419
Significance F	0.0293	0.1374

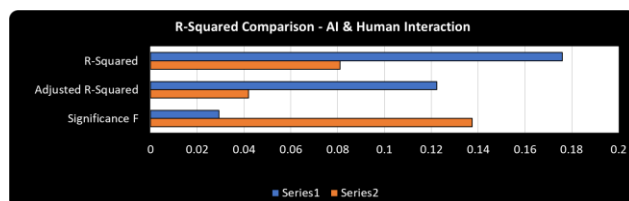


Figure 2 – Summary of Model Fit Metrics (Source: Author)

Table 4, Figures 3–4 details predictors, confirming consistency as the sole significant factor for AI ($p < 0.05$). The coefficient for

efficiency (Q1_a) in AI interactions is positive, but its p -value (0.0583) is slightly above the conventional significance threshold, indicating a weak but positive relationship with satisfaction. Consistency (Q1_c) in AI interactions is statistically significant ($p = 0.0351$) and represents a strong driver of satisfaction.

Response time (Q5_a) metrics for AI are not significant ($p = 0.2716$), suggesting that response time might not strongly influence AI satisfaction levels. Regarding human interactions, the coefficient for personalization (Q1_b) is negative and not significant ($p = 0.2161$), implying that personalization may not strongly predict satisfaction. Empathy (Q1_d), however, shows marginal significance ($p = 0.0539$), indicating that empathy plays a somewhat important role in satisfaction with human interactions.

Table 4 – Predictor Analysis (Source: Author)

Factor	Model	Coefficient	P-Value	Significant?
Efficiency (Q1_a)	AI	0.1827	0.0583	No
Consistency (Q1_c)	AI	0.1762	0.0351	Yes
Response Time (Q5_a)	AI	0.1569	0.2716	No
Personalization (Q1_b)	Human	-0.1481	0.2161	No
Empathy (Q1_d)	Human	0.2479	0.0539	Marginal

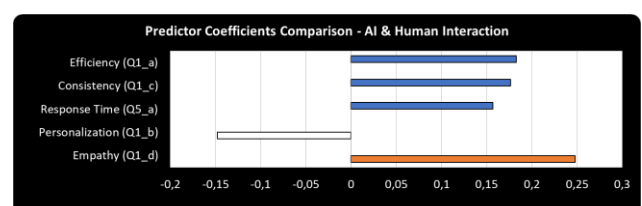


Figure 3 – Predictor Coefficient Comparison (Source: Author)

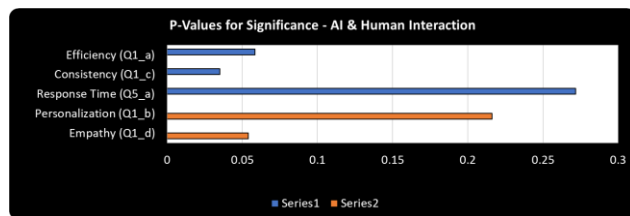


Figure 4 – P-Values for Significance (Source: Author)

For the AI satisfaction metric (Table 5), the F -value of 3.2762 with a p -value of 0.0293 confirms that the model is statistically significant, indicating that the predictors—efficiency, consistency, and response time—collectively explain satisfaction well. In contrast, for the human interaction metric, the F -value of 2.0714 with a p -value of 0.1374 suggests that the predictors—personalization and empathy—are not strong enough to explain satisfaction significantly.

Table 5 –Detailed ANOVA Table (Source: Author)

Model	df	SS	MS	F-Value	Significance F
AI	3	3.9294	1.31	3.2762	0.0293
	46	18.391	0.4		
Total	49	22.32			
Human	2	2.1077	1.054	2.0714	0.1374
	47	23.912	0.509		
Total	49	26.02			

Table 6 – Regression for AI Preferences (Source: Author)

Regression Statistics	
Multiple R	0.20336433
R Square	0.041357051
Adjusted R Square	-0.021163142
Standard Error	0.682019231
Observations	50

ANOVA

	df	SS	MS	F	Significance F
Regression	3	0.923089374	0.307696458	0.661499097	0.579967056
Residual	46	21.39691063	0.465150231		
Total	49	22.32			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	2.583344052	0.794889478	3.249941186	0.002161458	0.983314521	4.183373584	0.983314521	4.183373584
Q4 d - Empathy & Understanding	0.002391384	0.116565658	0.020515336	0.98372101	-0.232243117	0.237025884	-0.232243117	0.237025884
Q5 a - AI Response Time	0.136346932	0.155225877	0.878377591	0.384304004	-0.176106552	0.448800416	-0.176106552	0.448800416
Q1 b - Human Personalization	0.120415749	0.103819976	1.159851442	0.252094134	-0.088563024	0.329394522	-0.088563024	0.329394522

Key Takeaways. AI satisfaction is likely influenced by factors beyond empathy, response time, and personalization. Elements such as accuracy, reliability, or even trust in the AI system could be more significant in determining customer satisfaction.

Another regression analysis was conducted to examine how human empathy, convenience and speed, personalized service, and empathy and understanding influence human satisfaction in the U.S. retail banking system. A few key points from that analysis are presented below. Table 7 presents the results indicating $R^2 = 0.220$, $F(4,45) = 3.18$, and $p = 0.022$, with personalization found to be statistically significant ($\beta = 0.316$, $p = 0.010$). The R^2 value of 22.02% demonstrates that the model explains about 22% of the variance in human satisfaction, suggesting that certain unaccounted factors may also influence this outcome. The significance value $F = 0.0221$ confirms that the overall model is statistically significant, implying that the factors included are meaningfully related to satisfaction levels.

Key Takeaways. Human touch plays a notably significant role in determining human satisfaction. In contrast, human empathy as well as empathy and understanding exhibit weak and statistically insignificant effects, indicating that they are not critical within this model. Similarly, convenience and speed do not appear to exert a significant impact on satisfaction.

Comparative Analysis of AI and Human Interaction – Customer Experiences. To achieve a clearer understanding of the

Research Question 3: Customer Preferences

The purpose of this analysis is to understand customer preferences between AI and human interaction in retail banking branches, focusing on empathy, understanding, response time, and human personalization. A regression analysis was conducted to explore the influence of empathy, response time, and personalization on AI satisfaction. According to Table 6 (Regression for AI Preferences), the results show $R^2 = 0.041$ and $F(3,46) = 0.66$ with $p = 0.580$, indicating weak predictive power.

The R^2 value of 4.14 % suggests that only about 4% of the variance in AI satisfaction is explained by the model, implying that other factors, such as trust in AI or user expectations, may play a more significant role. The model significance value of 0.58 confirms that it is not statistically significant, meaning empathy, response time, and personalization cannot strongly predict AI satisfaction. The empathy coefficient (0.0024) and its high p -value (0.984) further demonstrate that empathy has minimal impact within this dataset.

A key observation from this pilot study was the inconsistent effect of the personalization variable across different regression models (see Table 2 versus Table 7). Whereas in the initial model personalization was a non-significant predictor of human satisfaction, in the broader preferences model it became a significant positive predictor. This contradiction emphasizes the methodological complexity of measuring this construction and points to a critical area for refinement in further research. It suggests that the relationship between personalization and satisfaction can be highly sensitive to other variables in the model - a phenomenon that should be further investigated with a larger, more diverse sample.

comparison between AI and human interaction, a comparative analysis was carried out to identify which factors better predict customer experiences. Table 8, Figure 5 compares overall model fit, emphasizing the human model's stronger correlation (higher Multiple R and R^2). The Multiple R value for AI is 0.2034, which is lower than the Human Multiple R value of 0.4693, meaning human interactions have a stronger overall correlation with customer satisfaction. R -Square for AI is 0.0414, while for Human Interaction it is 0.2202, meaning the AI and Human models explain 4.14% and 22.02% of the variance in customer satisfaction, respectively. The Adjusted R -Square for AI is negative (-0.0212), suggesting that the AI model is unfit here, whereas the Human model has a positive Adjusted R -Square (0.1509).

Table 9, Figure 6 confirms the significance of the human model ($p = 0.022$), in contrast to the AI model ($p = 0.58$). The AI model has an F -value of 0.6615 with a Significance F of 0.5800, indicating that it is not statistically significant. In comparison, the human model shows an F -value of 3.1777 with a Significance F of 0.0221, which is statistically significant. This means that human interactions have a meaningful impact on customer satisfaction, whereas AI interactions do not.

Table 10, Figures 7–8 analyzes the regression coefficients, highlighting the importance of personalization in human satisfaction ($p < 0.01$). For the AI model, the intercept is statistically significant ($p = 0.0022$), while other factors—such as Empathy &

Understanding (Q4_d) and AI Response Time (Q5_a)—have non-significant p -values, indicating that they do not notably influence customer satisfaction. This may imply that satisfaction with AI

depends more on the correctness of outcomes than on emotional or temporal aspects.

Table 7: Regression for Human Preferences (Source: Author)

Regression Statistics	
Multiple R	0.469307789
R Square	0.220249801
Adjusted R Square	0.150938672
Standard Error	0.671467715
Observations	50

ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	4	5.730899828	1.432724957	3.177697508	0.022076267
Residual	45	20.28910017	0.450868893		
Total	49	26.02			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	2.977028355	0.842351138	3.53418927	0.000959684	1.280446072	4.673610637	1.280446072	4.673610637
Q1_d - Human Empathy	0.142870111	0.113817726	1.255253609	0.215866235	-0.086370556	0.372110778	-0.086370556	0.372110778
Q4_a - Convenience & Speed	-0.119646168	0.135803144	-0.881026493	0.382984977	-0.393167741	0.153875405	-0.393167741	0.153875405
Q4_b - Personalized Service	0.315858923	0.117419221	2.690010376	0.009989059	0.079364472	0.552353374	0.079364472	0.552353374
Q4_d - Empathy & Understanding	-0.064905441	0.118868068	-0.546029238	0.58774282	-0.30431802	0.174507138	-0.30431802	0.174507138

Table 8 – Summary of Model Fit Metrics (Source: Author)

Metric	AI Value	Human Value
Multiple R	0.2034	0.4693
R-Square	0.0414	0.2202
Adjusted R-Square	-0.0212	0.1509
Standard Error	0.682	0.6715
Observations	50	50

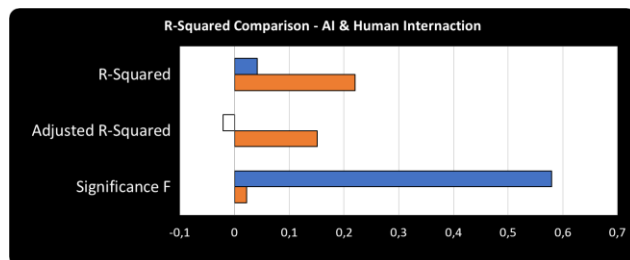


Figure 5 – R-Squared Comparison (Source: Author)

For the human model, Personalized Service (Q4_b) emerges as the only significant predictor of satisfaction, with a positive coefficient of 0.3159 and a p -value of 0.00999, showing a clear positive impact on satisfaction. In contrast, Human Empathy (Q1_d) and Convenience & Speed (Q4_a) are not statistically significant ($p = 0.2159$ and 0.3830 , respectively), suggesting that these factors exert minimal influence on customer satisfaction.

Table 9 – ANOVA Summary (Source: Author)

Model	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F-Value</i>	<i>Significance F</i>
AI	3	0.9231	0.308	0.6615	0.58
	46	21.397	0.465		
Total	49	22.32			
Model	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F-Value</i>	<i>Significance F</i>
Human	4	5.7309	1.433	3.1777	0.0221
	45	20.289	0.451		
Total	49	26.02			

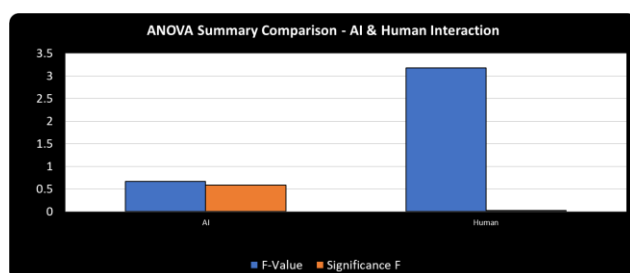


Figure 6 – ANOVA Summary Comparison (Source: Author)

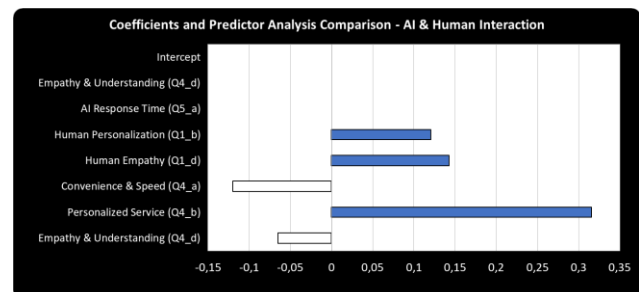


Figure 7 – Coefficients and Predictor Analysis Comparison (Source: Author)

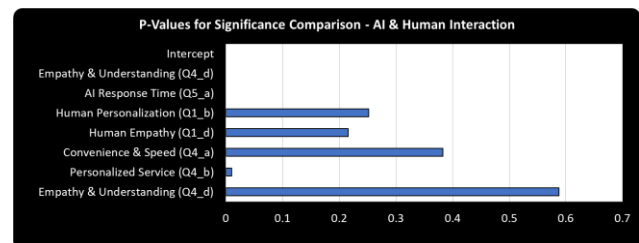


Figure 8 – P-Values for Significance Comparison (Source: Author)

Discussion

The assessment of customer satisfaction encompassed both AI and human interactions. Given that AI satisfaction consistently scores lower, further investigation into user preferences and the inherent limitations of AI-based interactions appears warranted, particularly within the context of the U.S. retail banking system. Additionally, considering demographic strategies—such as accounting for age and gender—may enhance the design of both AI systems and human-centered services.

These demographic patterns are consistent with findings reported by Méndez-Suárez et al. (2023), which support the implementation of segmented service strategies: AI solutions may be more effective for efficiency-driven younger male cohorts, whereas human interactions remain preferable for empathy-oriented female and older customer groups. Such differentiation could optimize economic outcomes, including customer loyalty.

Notably, younger males tend to rate AI performance higher, indicating greater openness or comfort with AI-driven technologies. Future studies should further examine the comparative quality of AI and human interactions, with particular focus on AI's conversational competence and its effectiveness in problem-solving.

Table 10 – Coefficients and Predictor Analysis (Source: Author)

Factor	Model	Coefficient	Standard Error	t Stat	P-Value	Yes
Intercept	AI	2.5833	0.7949	3.25	0.0022	No
Empathy & Understanding (Q4 d)	AI	0.0024	0.1166	0.021	0.9837	No
AI Response Time (Q5 a)	AI	0.1363	0.1552	0.878	0.3843	No
Human Personalization (Q1 b)	Human	0.1204	0.1038	1.16	0.2521	No
Human Empathy (Q1 d)	Human	0.1429	0.1138	1.255	0.2159	No
Convenience & Speed (Q4 a)	Human	-0.1196	0.1358	-0.881	0.383	Yes
Personalized Service (Q4 b)	Human	0.3159	0.1174	2.69	0.00999	No
Empathy & Understanding (Q4 d)	Human	-0.0649	0.1189	-0.546	0.5877	

Assessing Customer Feedback: AI or Human. Customer feedback was assessed in terms of satisfaction with AI and human interactions, focusing on efficiency, consistency, and response time, which together shape the overall experience. The analysis emphasizes the need to improve consistency in AI-driven technologies and services, as this factor has the strongest and most reliable impact on satisfaction. Attention should also be given to efficiency, since it represents an area with potential for improvement.

Comparative Analysis of AI and Human Interaction. In the comparative context, AI-driven technologies in the U.S. retail banking system should concentrate on improving efficiency and response time to enhance overall customer satisfaction. At the same time, organizations may invest more in employee training to strengthen empathy, which is approaching significance and could foster better development of human interaction-related services.

The regression results (Tables 1–2) underscore the role of consistency in AI models ($R^2 = 0.176$) compared with personalization in human models ($R^2 = 0.220$), supporting hybrid approaches (Huang & Rust, 2022) that balance cost efficiencies with relational value in banking economics.

Conclusions

Drawing on large-scale studies of U.S. retail banking, human contact continues to hold a significant advantage in delivering customer satisfaction over AI-based services by 0.58 points. This distinction is demographically sensitive, with a clear preference for human contact among women across all age ranges, with those aged 30–39 showing the strongest preference (mean = 4.24), whereas young men (22–29) are more tolerant of AI. Drivers, in turn, further elucidate this difference: AI effectiveness is most correlated with consistency, whereas human effectiveness is strongly correlated with individualized service, with empathy itself bordering on statistical significance. Economically, the far greater explanatory power of human interaction ($R^2 = 22\%$ versus 4% for AI) suggests that underlying forces such as trust and relationship building disproportionately influence customer loyalty. Thus, for policymakers and bankers the road ahead is twofold: optimizing AI for uniformity in automated systems and investing in human capital to provide hyper-personalized service, particularly for key demographic segments. Finally, as AI brings incremental efficiency benefits, the findings endorse that human connection remains the primary driver of increased satisfaction, implying a hybrid future—consistent with the "Caring Machine" concept—where technology enhances human emotional intelligence rather than replacing it to foster loyalty and economic security (Huang & Rust, 2022).

References

- Ameen, N., Tarhini, A., Shah, M. H., & Nusair, K. (2021). A cross cultural study of gender differences in omnichannel retailing contexts. *Journal of Retailing and Consumer Services*, 58, 102265. <https://doi.org/10.1016/j.jretconser.2020.102265>
- Calabrese, A., Costa, R., & Rosati, F. (2016). Gender differences in customer expectations and perceptions of corporate social responsibility. *Journal of Cleaner Production*, 116, 135–149. <https://doi.org/10.1016/j.jclepro.2015.12.100>
- Deloitte Network. (2024). *New Deloitte Survey: Increasing Consumer Privacy and Security Concerns in the Generative AI Era*. Deloitte Touche Tohmatsu Limited. <https://e.surl.li/xuzbdz>
- Huang, M.-H., & Rust, R. T. (2022). A Framework for Collaborative Artificial Intelligence in Marketing. *Journal of Retailing*, 98(2), 209–223. <https://doi.org/10.1016/j.jretai.2021.03.001>
- McKinsey & Company. (2021). *Building the AI bank of the future: Global Banking Practice*. McKinsey & Company. <https://e.surl.li/tnwpvw>
- Méndez-Suárez, M., Monfort, A., & Hervás-Oliver, J.-L. (2023). Are you adopting artificial intelligence products? Social-demographic factors to explain customer acceptance. *European Research on Management and Business Economics*, 29(3), 100223. <https://doi.org/10.1016/j.iemeen.2023.100223>
- Pattanayak, S. K. (2021). The Impact of Artificial Intelligence on Operational Efficiency in Banking: A Comprehensive Analysis of Automation and Process Optimization. *International Research Journal of Engineering and Technology*, 8(10), 2049–2061. <https://www.irjet.net/archives/V8/i10/IRJET-V8I10315.pdf>
- Scheffler, P., & Puczyk, A. (2025). *AI in Retail Banking: Use Cases, Challenges, and Trends*. Neontri. <https://neontri.com/blog/ai-retail-banking/>
- The Contentstack Team. (2024). *Personalized experiences: How emotional intelligence improves customer experience*. Contentstack Inc. <https://e.surl.li/lkguqv>

Limitations

First and foremost, being a pilot study, the greatest limitations lie in its scale and scope. A clear acknowledgement of these limits is important to avoid any potential misinterpretation of its results and to frame correctly the contribution this makes as a first step. The study has a low sample size, $n = 50$, exclusively drawn from a university population; thus, findings explicitly have limited generalizability. The use of convenience sampling introduces a risk of selection bias, while self-reported measures may be affected by recall errors. The cross-sectional design also precludes causal inference, and uncontrolled factors (e.g., prior AI exposure) may confound the results. Future studies should address these limitations through larger and more diverse samples, complemented by longitudinal research designs.

Future Research

The insights gained from this pilot study highlight a very clear direction in which further research should go. The next logical step involves the verification of these findings with at least a sample size of 500 participants, with diverse characteristics, using longitudinal measures that allow for tracking satisfaction over time. With a more robust foundation like that, future hybrid AI-human models could examine key critical variables such as trust and cultural influence. Meanwhile, advanced analytical techniques, such as structural equation modeling, could be utilized to unpack complex mediation effects, including, but not limited to, examining whether empathy acts as a mediator of the personalization and user satisfaction relationship. This multi-phase research agenda would go a long way in strengthening the applicability and impact of this study to U.S. banking innovation and policy.

Conflicts of Interest

Since this is a self-authored manuscript, no conflict of interest exists.

Informed Consent Declaration

This research involved human participants. As a pilot study, the population size was intentionally kept small ($n = 50$) to ensure feasibility. The participants were drawn from both graduate and undergraduate students at Wright State University.