# Understanding Customer Satisfaction of Chatbots Service and System Quality in Banking Services

**Suraya Zainol**

Ph.D Candidate, Universiti of Kuala Lumpur, (UniKL) Business School, Jalan Gurney, Kampung Datuk Keramat, 54000 Kuala Lumpur, Malaysia. E-mail: suraya.zainol08@s.unikl.edu.my

**Mohd Farid Shamsudin\*** ⓘD

\*Corresponding author, Associate Prof., University of Kuala Lumpur, (UniKL) Business School, Jalan Gurney, Kampung Datuk Keramat, 54000 Kuala Lumpur, Malaysia. E-mail: mfarid@unikl.edu.my

**Sallaudin Hassan**

University Kuala Lumpur (MITEC) Jalan Persiaran Sinaran Ilmu, Bandar Seri Alam. 81750 Masai, Johor. E-mail: sallaudin@unikl.edu.my

**Nor Azila Mohd Noor** ⓘD

University Utara Malaysia, School of Business Management (SBM), 06010 Sintok, Kedah, Malaysia. Email: azila@uum.edu.my

## Abstract

Chatbots is a computer software powered by artificial intelligence designed to replicate human interaction. It is also possible to refer to them as digital assistants that comprehend the capacities of humans. The bot interprets the user's intent, then processes their queries and provides prompt responses. Chatbots perform their most crucial role: to analyse and detect the intent of the user's request to extract relevant entities. AI-powered chatbots were introduced to improve operational efficiency, eventually saving organisational costs. This study investigates the role of system and service quality in customer satisfaction in banking services. One hundred forty-five usable data were used for analysis. Data were analysed using the Smart PLS. The results revealed that response time, usability, adaptability, empathy and responsiveness were insignificant for customer satisfaction. The result is important as it gave the insight point of customers with regards to the new services. Business organisations may need to introduce chatbots and perhaps make some improvements from time to time to provide better services.

# Introduction

Today, an Ai-based messaging platform called a chatbot has been introduced to solve actual issues. The chatbot is designed to act like conversational agents with voice-enabled, text-based, or a combination of both  (Jenneboer et al., 2022). Text-based devices can read and respond to written messages and requests, while voice-enabled devices can listen to and respond to spoken messages and requests, such as apple's Siri and amazon's Alexa. All conversational agents rely on a series of underlying core technologies to understand natural language information and human intention and thus engage in human-like conversational interactions (Cheng & Jiang, 2020).

Chatbots in retail, e-commerce, banks, and health services have been explored to save more than $8 billion per year by 2022, up from $20 million this year (Foye, 2017). A chatbot is one of artificial intelligence's innovative and rapid growth fields, with a market estimate of $1.3 billion by 2025 (Juniper, 2017). It is expected to save USD 8 billion per annum by 2022. The global chatbot market is expected to expand from US$2.6 billion to US$9.4 billion between 2019 and 2024, further than massively increasing in five years (FMI, 2022). The Asia Pacific chatbot market is expanding at a 9% compound annual growth rate (CAGR). The demand in North America is rising at a 31.2% CAGR, while Europe is growing at 30.4% (Juniper, 2017). The chatbot is actively used in five countries: the United States, India, Germany, the United Kingdom, and Brazil. Chatbot participation has risen fivefold in the last five years (Beckmann, 2022).

# Literature Review

## Customer satisfaction

At the most basic, customer satisfaction is achievable when a business's services satisfy customer expectations (Eren, 2021). Consumers and service providers frequently view excellent service as a sign of quality. Thus it requires effort to enhance the quality of the service system offered to meet and exceed user expectations and satisfaction (Nguyen et al., 2020). As a result, the organisation must place a premium on service quality to ensure client happiness (Djunaidi, Tenriawali & Umanailo, 2020). Chatbots were designed to provide a high level of services to customers. Today, the competition among service providers is getting

tough. Industry players compete with each other to attract more customers and increase their market share. According to Ong et al. (2017), customer satisfaction is important in leading to sales revenue growth. Besides that, customer satisfaction also helps to contribute to increasing brand popularity (Braxton & Lau-Gesk, 2020). Indirectly, customer satisfaction helps service providers during a corporate crisis. In addition, it is reported that customer satisfaction helps customers to stand out from competitors (Kataria & Saini, 2019).

## System Quality

Abror et al. (2020) discovered that system quality should be prioritised. Since a chatbot is a system component, maintaining the system's functionality is advantageous to maintain the system's functionality. This can be accomplished through routine system maintenance and installing a backup system capable of temporarily replacing the primary system in the case of system failures or obstacles acting on the primary system (Jenneboer et al., 2022). Furthermore, as the features supplied by chatbots grow, sustaining the most recent version becomes vital. Geoghegan et al. (2021) highlighted that businesses should avoid building chatbots on platforms not currently supported by other systems. Chatbot service requires a system that is closely tied to platform appropriateness. In theory, any web-based platform can construct a chatbot, but the media via which it can be accessed is limited (Mohamad Suhaili et al., 2021). Ilievski et al. (2018) recommended that businesses avoid employing chatbots with a platform not widely supported by other systems.

Gao and Jiang (2021) stated in their study that system quality is determined by the app's usability, availability, dependability, adaptability, and reaction time. Users often prefer to complete the intended action with the fewest possible clicks, suggesting a propensity for apps that require the fewest possible clicks to complete the operation (Denecke & Warren, 2020). Additionally, users value the app's availability and ability to deliver accessible, consistent, rapid, and appropriate operational performance. This demonstrates that app enjoyment may suffer if an app does not function as is expected when required.

## Service quality

Service quality refers to the information system provider's responsiveness, assurance, and empathy (Yao et al., 2020). According to Kant and Jaiswal (2017), end-user computing forced information systems to adopt two roles: one as a source of information and another as a provider of services. They established that the quality of service also contributed to the success or failure of an information system.

Yilmaz et al. (2018) defined reliability as an information system's capacity to operate properly over a specified time and under various operational situations. Sometimes, a consumer may experience a slow internet connection or be in an area with weak network signals. Despite this, consumers want the app to work effectively, which presents a hurdle. Responsiveness relates to a business's ability to provide service to its customers promptly (Teeroovengadum, 2020). At times, mobile app backend operations support may be

insufficient, resulting in discontent. Assurance was described as the ability of employees to carry out their given jobs (Mulia et al., 2020). For example, most internet businesses operate a call centre for customer care. Suppose a customer support professional is not adequately trained. In that case, the consumer will soon spread the app news and quit using it. Ahmed et al. (2017) defined empathy as valuing the user's interests. A consumer, for example, may be uninformed of a business's loyalty programme. Following that, the organisation needs to engage with clients and educate them about the multiple benefits offered by their loyalty programme. This greeting tells the consumer that the firm is looking for her best interests (Lau et al., 2019).

## Methodology

A questionnaire was constructed to investigate the study hypotheses that were provided. The questionnaire was distributed to roughly 700 individuals using various messaging and social networking tools (Instagram, LinkedIn and WhatsApp). It was required of the participants that they distribute the questionnaire to their network. As a result, a combination of convenience and snowball sampling methodologies was applied. Participants in the study were required to be at least 18 years old and have previous experience using chatbot services provided by one banking institution in Malaysia. The questionnaire was filled out between October 2021 and December 2021. However, because chatbots are still relatively new to the banking industry, the number of consumers who could be identified as having prior experience with a bank's chatbot service was limited to 156. Twenty-one of these gave incomplete responses. As a result, the final sample used in the statistical analysis consisted of only 145 cases. 35% of the participants are women (51), 60% of the participants are married (87), 70% of the participants are between the ages of 30 and 40 (101), and 80% of the participants hold undergraduate degrees (116).

Before commencing the survey, participants were given a basic introduction to chatbots. We solicited feedback from participants regarding the effectiveness of the bank's chatbot-based customer support (operating via text and voice). The initial part of the research involved us investigating the customer's prior experiences with chatbots and their expectations of those services. In the second section, we concentrated on the demographic aspects of the people who took part in the study. The items included in this study's scales were derived from previous research. On a scale similar to the Likert, each item was given a score between one and seven, with one representing "Strongly disagree" and seven representing "Strongly agree" (7). A questionnaire that contained these items underwent preliminary testing. Additionally, it was thoroughly evaluated and finalised by specialists from the banking sector and the academic community.

# Findings

Table 1: Confirmatory Factor Analysis Results

| Constructs | No of the items used | No. of items deleted | Items retained |
|---|---|---|---|
| Customer Satisfaction | 4 | 0 | 4 |
| | Csat1 | | Csat1 |
| | Csat2 | | Csat2 |
| | Csat3 | | Csat3 |
| | Csat4 | | Csat4 |
| Response time | 5 | 2 | 3 |
| | ResT1 | | ResT1 |
| | ResT2 | ResT2 | |
| | ResT3 | | ResT3 |
| | ResT4 | ResT4 | |
| | ResT5 | | ResT5 |
| Usability | 4 | 1 | 3 |
| | Use1 | | Use1 |
| | Use2 | Use2 | |
| | Use3 | | Use3 |
| | Use4 | | Use4 |
| Reliability | 3 | 0 | 3 |
| | Reliab1 | | Reliab1 |
| | Reliab2 | | Reliab2 |
| | Reliab 3 | | Reliab3 |
| Availability | 3 | 0 | 3 |
| Avail1 | Avail1 | | Avail1 |
| Avail2 | Avail2 | | Avail2 |
| Avail3 | Avail3 | | Avail3 |
| Adaptability | 6 | 3 | 3 |
| | Adapt1 | Adapt1 | |
| | Adapt2 | Adapt2 | |
| | Adapt3 | | Adapt3 |
| | Adapt4 | | Adapt4 |
| | Adapt5 | | Adapt5 |
| | Adapt6 | Adapt6 | |
| Service Quality | | | |
| Assurance | 4 | 1 | 3 |
| Assu1 | Assu1 | | Assu1 |
| Assu2 | Assu2 | | Assu2 |
| Assu3 | Assu3 | Assu3 | |
| Assu4 | Assu4 | | Assu4 |
| Empathy | 3 | 0 | 3 |
| | Em1 | | Em1 |
| | Em2 | | Em2 |
| | Em3 | | Em3 |
| Responsiveness | 3 | 1 | 3 |
| | Res1 | | Res1 |
| | Res2 | | Res2 |
| | Res3 | | Res3 |
| | Res4 | Res4 | |

Table 1 indicates the items per construct used in the study, with the details of items deleted and retained for the next analysis step. Based on Table 1, RestT2, Rest4, Use2, Adapt1, Adapt2, Adapt6, Assu3 and Res4 were deleted due to low value ( <0.7). According to Hair et al. (2012), deletion of items should only be done if the indicator's reliability is low. Such deletion will significantly increase AVE and CR.

Table 2: Internal Consistency, Indicator Reliability, and Convergent Validity Test

| Indicator | Construct | Outer loadings (> 0.7) | AVE (>0.5) | CR (>0.7) | Cronbach's Alpha (>0.7) |
|---|---|---|---|---|---|
| Csat1 | Customer Satisfaction | 0.788 | 0.699 | 0.867 | 0.87 |
| Csat2 | | 0.763 | | | |
| Csat3 | | 0.771 | | | |
| Csat4 | | 0.841 | | | |
| ResT1 | Response time | 0.789 | 0.669 | 0.934 | 0.871 |
| ResT3 | | 0.835 | | | |
| ResT5 | | 0.797 | | | |
| Use1 | Usability | 0.741 | 0.566 | 0.839 | 0.750 |
| Use3 | | 0.792 | | | |
| Use4 | | 0.865 | | | |
| Reliab1 | Reliability | 0.780 | 0.649 | 0.880 | 0.766 |
| Reliab2 | | 0.799 | | | |
| Reliab 3 | | 0.836 | | | |
| Avail1 | Availability | 0.860 | 0.678 | 0.824 | 0.733 |
| Avail2 | | 0.901 | | | |
| Avail3 | | 0.779 | | | |
| Adapt3 | Adaptability | 0.849 | 0.547 | 0.828 | 0.826 |
| Adapt4 | | 0.723 | | | |
| Adapt5 | | 0.832 | | | |
| Assu1 | Assurance | 0.822 | 0.653 | 0.849 | 0.831 |
| Assu2 | | 0.848 | | | |
| Assu4 | | 0.892 | | | |
| Em1 | Empathy | 0.823 | 0.721 | 0.912 | 0.795 |
| Em2 | | 0.749 | | | |
| Em3 | | 0.817 | | | |
| Res1 | Responsiveness | 0.726 | 0.682 | 0.937 | 0.846 |
| Res2 | | 0.752 | | | |
| Res3 | | 0.719 | | | |

Table 2 displays the items used with the AVE, CR and Cronbach's Alpha value. All items display acceptable value within the rule of thumb based on the suggestion by Hair & Fávero (2019).

Table 3: Discriminant Validity using Fornell-Larcker Criterion

| | Customer Satisfaction | Response time | Usability | Reliability | Availability | Assurance | Empathy | Responsiveness |
|---|---|---|---|---|---|---|---|---|
| Customer Satisfaction | 0.889 | | | | | | | |
| Response time | 0.879 | 0.824 | | | | | | |
| Usability | 0.826 | 0.818 | | | | | | |
| Reliability | 0.739 | 0.526 | 0.852 | | | | | |
| Availability | 0.742 | 0.439 | 0.370 | 0.849 | | | | |
| Adaptability | 0.666 | 0.242 | 0.364 | 0.594 | 0.790 | | | |
| Assurance | 0.574 | 0.466 | 0.638 | 0.340 | 0.401 | 0.808 | | |
| Empathy | 0.561 | 0.547 | 0.460 | 0.375 | 0.352 | 0.465 | 0.740 | |
| Responsiveness | 0.421 | 0.560 | 0.387 | 0.632 | 0.475 | 0.368 | 0.366 | 0.805 |

The above table indicates the criterion developed by Fornell and Larcker (1981). The purpose is to determine the discriminant validity of a test. The statistics reveal that the square root of each construct's diagonal value should be higher than the rest in the row. The results show no problem with the discriminant validity of the study and the items used.

Table 4: Discriminant Validity using HTMT Approach

| | Customer Satisfaction | Response time | Usability | Reliability | Availability | Assurance | Empathy | Responsiveness |
|---|---|---|---|---|---|---|---|---|
| Customer Satisfaction | | | | | | | | |
| Response time | 0.879 | | | | | | | |
| Usability | 0.826 | 0.818 | | | | | | |
| Reliability | 0.739 | 0.526 | 0.852 | | | | | |
| Availability | 0.742 | 0.439 | 0.370 | 0.849 | | | | |
| Adaptability | 0.666 | 0.242 | 0.364 | 0.594 | 0.790 | | | |
| Assurance | 0.574 | 0.466 | 0.638 | 0.340 | 0.401 | 0.808 | | |
| Empathy | 0.561 | 0.547 | 0.460 | 0.375 | 0.352 | 0.465 | 0.740 | |
| Responsiveness | 0.421 | 0.560 | 0.387 | 0.632 | 0.475 | 0.368 | 0.366 | |

The HTMT result reveals that all values are lower than the minimum cutoff. The conservative criterion shows discriminant validity when the HTMT is 90%. The result revealed no discriminant validity problem because all constructs had below 0.90 values. In other words, discriminant validity is demonstrated in this study.

Table 5: Direct relationship measurement

| | Original Sample (O) | Sample Mean (M) | Standard Deviation (STDEV) | T Statistics (\|O/STDEV\|) | P Values |
|---|---|---|---|---|---|
| Response time → Customer Satisfaction | 0.055 | 0.057 | 0.057 | 0.954 | 0.340 |
| Usability → Customer Satisfaction | 0.078 | 0.074 | 0.068 | 0.855 | 0.481 |
| Reliability → Customer Satisfaction | 0.493 | 0.486 | 0.053 | 9.261 | 0.000 |
| Availability → Customer Satisfaction | 0.482 | 0.474 | 0.057 | 8.493 | 0.000 |
| Adaptability → Customer Satisfaction | 0.054 | 0.051 | 0.035 | 1.556 | 0.120 |
| Assurance → Customer Satisfaction | 0.100 | 0.097 | 0.040 | 2.514 | 0.012 |
| Empathy → Customer Satisfaction | 0.037 | 0.042 | 0.026 | 1.413 | 0.158 |
| Responsiveness → Customer Satisfaction | 0.022 | 0.021 | 0.015 | 1.522 | 0.129 |

The result of path coefficients indicates that only two factors from the system quality significantly impacted customer satisfaction. Reliability and availability positively impacted customer satisfaction regarding the chatbot's experience. The result revealed that response time, usability and adaptability were insignificant to customer satisfaction.

On the other part, only assurance was positively significant towards customer satisfaction. Empathy and responsiveness were insignificant. Such results revealed are interesting and will be discussed in the next section.

## Discussions and Conclusions

The results revealed that response time, usability and adaptability were insignificant for customer satisfaction. Being a new system, chatbots could be perceived as a new technology that complicates and changes customers' norms. People are reluctant to change and adapt to a new system. The way chatbots work could not be similar to how normal humans render services. Therefore, it could take time for customers to adapt to the technology. Although chatbots have been programmed and designed to solve basic questions, it may take some time for customers to get familiarised. New technology perhaps may need some time to be perfect. Response time could not be as perfect as humans who can use their judgements and personal experiences to handle situations. At the same time, chatbots rely on the program that has been installed. Chatbots need to go through a process that chatbots need to go through that may delay the response time. In the meantime, customers should take the opportunity and look at the positive sides of the system's ability to solve common and basic questions or enquiries.

From the perspective of service quality, empathy and responsiveness were found to be insignificant to customer satisfaction. Based on the result, it can be concluded that customers perceived the chatbots as the perfect solution, which leads to poor feedback. Five insignificant

constructs can be easily associated with those closely related to humans. Response time, usability and adaptability are closely related to human matters. High expectations from customers on the chatbot's performance could lead to frustration. As a result, customers expect chatbots to respond and show some empathy as what they have enjoyed in the past.

Chatbots in Malaysia can be considered new in which there is a need for all relevant parties to educate and create some awareness of them. In the near future, chatbots could be accepted as a more new programs are installed and designed to meet customer expectations.

## Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article

# References

Abror, A., Patrisia, D., Engriani, Y., Evanita, S., Yasri, Y., & Dastgir, S. (2020). Service quality, religiosity, customer satisfaction, customer engagement and Islamic bank's customer loyalty. *Journal of Islamic Marketing*, *11*(6), 1691–1705. https://doi.org/10.1108/JIMA-03-2019-0044

Ahmed, S., Tarique, K. M., & Arif, I. (2017). Service quality, patient satisfaction and loyalty in the Bangladesh healthcare sector. *International Journal of Health Care Quality Assurance*, *30*(5), 477–488. https://doi.org/10.1108/IJHCQA-01-2017-0004

Braxton, D., & Lau-Gesk, L. (2020). The impact of collective brand personification on happiness and brand loyalty. *European Journal of Marketing*, *54*(10), 2365–2386. https://doi.org/10.1108/EJM-12-2019-0940

Cheng, Y., & Jiang, H. (2020). How Do AI-driven Chatbots Impact User Experience? Examining Gratifications, Perceived Privacy Risk, Satisfaction, Loyalty, and Continued Use. *Journal of Broadcasting and Electronic Media*, *64*(4), 592–614. https://doi.org/10.1080/08838151.2020.1834296

Denecke, K., & Warren, J. (2020). How to evaluate health applications with conversational user interface? *Studies in Health Technology and Informatics*. https://doi.org/10.3233/SHTI200307

Eren, B. A. (2021). Determinants of customer satisfaction in chatbot use: evidence from a banking application in Turkey. *International Journal of Bank Marketing*, *39*(2), 294–311. https://doi.org/10.1108/IJBM-02-2020-0056

Futuremarketinsight. (2022, August 17). *Chatbot Market*. Www.futuremarketinsights.com. https://www.futuremarketinsights.com/reports/chatbot-market?utm_source=adwords&utm_medium=ppc&utm_source=adwords&utm_medium=ppc&id=5&gclid=Cj0KCQjwgO2XBhCaARIsANrW2X0M3NqvmpelExG3m4QuApCBaljJMCboIjqR3vvwFaRT83jkLHb4tHAaAj8EEALw_wcB

Gao, Z., & Jiang, J. (2021). Evaluating Human-AI Hybrid Conversational Systems with Chatbot Message Suggestions. *International Conference on Information and Knowledge Management, Proceedings*. https://doi.org/10.1145/3459637.3482340

Hair, J. F., & Fávero, L. P. (2019). Multilevel modeling for longitudinal data: concepts and applications. *RAUSP Management Journal*, *54*(4), 459–489. https://doi.org/10.1108/RAUSP-04-2019-0059

Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The Use of Partial Least Squares Structural Equation Modeling in Strategic Management Research: A Review of Past Practices and Recommendations for Future Applications. *Long Range Planning*, *45*(5–6), 320–340. https://doi.org/10.1016/j.lrp.2012.09.008

Ilievski, V., Musat, C., Hossmann, A., & Baeriswyl, M. (2018). Goal-Oriented chatbot dialog management bootstrapping with transfer learning. *IJCAI International Joint Conference on Artificial Intelligence*. https://doi.org/10.24963/ijcai.2018/572

Jenneboer, L., Herrando, C., & Constantinides, E. (2022). The Impact of Chatbots on Customer Loyalty: A Systematic Literature Review. In *Journal of Theoretical and Applied Electronic Commerce Research*. https://doi.org/10.3390/jtaer17010011

Juniper. (2017, July 24). *Chatbot Conversations to deliver $8 billion in Cost savings by 2022*. Www.juniperresearch.com. https://www.juniperresearch.com/resources/analystxpress/july-2017/chatbot-conversations-to-deliver-8bn-cost-saving

Kant, R., & Jaiswal, D. (2017). The impact of perceived service quality dimensions on customer satisfaction: An empirical study on public sector banks in India. *International Journal of Bank*

*Marketing*, *35*(3), 411–430. https://doi.org/10.1108/IJBM-04-2016-0051

Kataria, S., & Saini, V. (2019). The mediating impact of customer satisfaction in relation of brand equity and brand loyalty: An empirical synthesis and re-examination. *South Asian Journal of Business Studies*, *9*(1), 62–87. https://doi.org/10.1108/SAJBS-03-2019-0046

Lau, T., Cheung, M. L., Pires, G. D., & Chan, C. (2019). Customer satisfaction with sommelier services of upscale Chinese restaurants in Hong Kong. *International Journal of Wine Business Research*, *ahead-of-p*(ahead-of-print). https://doi.org/10.1108/ijwbr-10-2018-0060

Mohamad Suhaili, S., Salim, N., & Jambli, M. N. (2021). Service chatbots: A systematic review. In *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2021.115461

Mulia, D., Usman, H., & Parwanto, N. B. (2020). The role of customer intimacy in increasing Islamic bank customer loyalty in using e-banking and m-banking. *Journal of Islamic Marketing*, *12*(6), 1097–1123. https://doi.org/10.1108/JIMA-09-2019-0190

Nguyen, X., Tran, H. L., Phan, H. A., & Phan, T. T. H. (2020). Factors influencing customer satisfaction: The case of facebook chabot vietnam. *International Journal of Data and Network Science*, *4*(2), 167–178. https://doi.org/10.5267/j.ijdns.2020.2.001

Ong, K. S., Nguyen, B., & Syed Alwi, S. F. (2017). Consumer-based virtual brand personality (CBVBP), customer satisfaction and brand loyalty in the online banking industry. *International Journal of Bank Marketing*, *35*(3), 370–390. https://doi.org/10.1108/IJBM-04-2016-0054

Teeroovengadum, V. (2020). Service quality dimensions as predictors of customer satisfaction and loyalty in the banking industry: moderating effects of gender. *European Business Review*. https://doi.org/10.1108/EBR-10-2019-0270

Yao, R., Song, S., Li, Q., Wang, C., Chen, H., Chen, H., & Zeng, D. D. (2020). Session-level user satisfaction prediction for customer service chatbot in E-commerce. *AAAI 2020 - 34th AAAI Conference on Artificial Intelligence*.

Yilmaz, V., Ari, E., & Gürbüz, H. (2018). Investigating the relationship between service quality dimensions, customer satisfaction and loyalty in Turkish banking sector: An application of structural equation model. *International Journal of Bank Marketing*, *36*(3), 423–440. https://doi.org/10.1108/IJBM-02-2017-0037