Balancing Automation and Human Support: The Impact of Chatbot Accuracy on Customer Trust in SMEs

Michon Goddijn[1]

[1]*Breda University Of Applied Sciences*

This manuscript was compiled on October 18, 2024

Abstract

In this study, the influence of chatbot accuracy on customer trust in comparison to traditional human support among small and medium sized enterprises is investigated to optimise customer service strategies. This paper employed a qualitative approach, in which semi structured interviews with participants who had interacted with both chatbots and human agents were analysed using thematic analysis. The literature review shows that chatbot reliability, the presence of a human fallback, and the congruency of chatbot responses and customer expectations all influence trust. Purposive sampling was used to collect data across different industries to capture a variety of experiences, and data were securely and systematically collected and analysed. Results highlight core themes of trust such as accuracy is necessary, human support is critical in chat bot limitations, customer frustration with inaccurate or generic responses, and the trade off between speed of response and personalization. The findings provide an indication that automating chatbots that intuitively answer customers' first questions is critical, but that is not the whole story, it can improve customer satisfaction by using accurate chatbots accompanied by conversations that escalate to human agents. Theoretical and practical contributions are made in this research that advocate for the strategic integration of technology and human support to engender customer trust and satisfaction that holds over time.

Keywords: *Chatbot accuracy, Customer trust, SMEs (Small and Medium Enterprises), Human fallback support, Customer satisfaction*

Corresponding author: Michon Goddijn, supervised by Tsegaye Tashu *E-mail address:* 231849@buas.nl.

1. Introduction

The use of automation in customer service is becoming more common as we use chatbots. For small and medium sized enterprises (SMEs), chatbots are a great thing to leverage as a way to provide instant response, lower costs and have an always available customer service. While this might be the case, customer trust is a decisive factor in the success of the chatbot implementation. Customers expect the chatbot to perform just like the customer service representative does. This study focuses on how the accuracy of chatbots in SMEs influence customer trust in comparison to conventional human assistance. This study aims to answer the research question: *How does the accuracy of chatbots in SMEs compare to traditional human assistance in terms of customer trust?* Studying the nuanced experiences of customers that interacted with human agents and chatbots, the paper examines how to optimise customer service strategies for SMEs. In this exploration, it digs into the importance of chatbot accuracy and customer's expectations, and how human fallback can be used to better meet them, as well as offer actionable strategies to improve customer satisfaction and trust in SMEs.

Strong relationships and customer loyalty in a competitive business environment depends on customer trust. If we think of trust as the relationship between an organization and a customer, then the customers will decide if they want to engage with that company, recommend its services and products, and to stay with them. In addition, this study also underscores the challenges faced by SMEs in successfully implementing chatbot technology. Unlike large corporations with huge resources, SMEs are constrained, so their chatbot solutions need to be accurate, reliable and trusted by their customers. In this research, we explore the conditions under which chatbots may either erode or build trust in customer service interactions by means of qualitative interviews.

2. Literature Review

SMEs are beginning to adopt chatbots for improving customer service. Araujo (2018) demonstrates that chatbot accuracy is related to customer trust in the chatbot and in the business. Trust is built with accurate responses and damaged by errors (Hill et al., 2015). Chatbots are often compared to human agents, which are typically perceived as better at handling complex questions because human agents can better provide personalised responses (Verhagen et al., 2014). This comparison drew attention to the gap in capabilities between chatbot and human agent and the need to bridge the gap in specifically complex or emotionally sensitive customer service situations.

Trustful chatbots are also a function of the chatbots design and functionality; for example, human-like features and empathy responses from chatbots increase customer trust in chatbots (Go & Sundar, 2019). And these features help to create a feeling of connexion and reliability that is often absent in automated interactions. Yet Diederich et al. (2019) claim a lack of contextual understanding or empathy can inhibit the growth of trust. To put it plainly, this is especially important when chatbots are seeking to answer inquiries beyond a surface level or that require an understanding of the customer's needs and feelings. Additionally, Shahid et al. (2020) argue that chatbots can still maintain trust by resolving an issue seamlessly and transferring the customer to a human agent when they fail. The trust in chatbot accuracy comes at a cost of human backup, but this balance is essential to not only maintaining a good customer experience but also to rebuild trust between chatbots and customers. Nevertheless, according to the literature, it is important to design the SME chatbot systems carefully; that is, to incorporate technological as well as human components that are conducive to smooth service transitions and well, provide personalized customer support.

There is also a large body of literature on the role of expectations in shaping customer's perception of chatbot interaction, Oliver (2014) explains that customer's expectation about chatbot is shaped by prior experiences with technology, marketing communication and cultural context. Customer trust is increased when chatbots meet or exceed these expectations. However, trust is eroded if chatbots fail to reach the anticipated level of performance, particularly in terms of accuracy, and customers cease to use the service again (Gefen & Straub, 2004). As this shows, this is whatever business needs to know businesses are going to need businesses to set realistic expectations with clear communication about chatbot capabilities and the limitations.

Further, Adam et al. (2020) show that the use of artificial intelligence (AI) and natural language processing (NLP) contributes to chatbot accuracy and thus better user experience. Chatbots with AI power up can learn from interactions, becoming more and more accurate with time. But the complexity of human language, and the complexities of meaning still pose persistent difficulties for chatbots, especially in emotional situations. The literature is unanimous in emphasising the need of human fallback support as a component of the chatbot implementation, so users get what they need when the technology is not up to the mark.

3. Methodology

This study used a qualitative research design to study customer experiences and perceptions of chatbot accuracy compared to traditional human customer service in SMEs. Participants who interacted with chatbots and human agents were interviewed on a semi structured basis.

3.1. Data Collection Methods

Initially, we planned to interview eight participants, which was reduced to five due to constraints on the number of participants available and redundancy of emerging themes during the first interviews. By the fifth interview, themes started to repeat themselves and the decision was made to prioritise depth over quantity. A smaller sample allowed a deeper dive into the individual experience while still capturing a good range of experiences across many different industries. The semi structure allowed exploration of participants' unique experiences, while key topics on chatbot accuracy and trust were covered. Sample questions included: "Can you share an experience where a chatbot gave you either correct or incorrect information?" and "How did the chatbot's accuracy affect your view of the company's service?" These questions were flexible enough so that the participants could give detailed, personal insights in the data collection process, giving the data a rich understanding of the issues.

The interviews were completed in a two week period as the participants' schedules dictated and the data collected was rich and representative of many customer experiences. The interviews took about 10 to 20 minutes each so there was plenty of time for the participants to talk about their experience and give thoughtful answers. To probe further into the participants' perceptions and feelings about interactions with chatbots and human agents, follow up questions were used.

3.2. Sampling Procedures

Purposive sampling was used to select participants who had previous experience with chatbots and human agents in SMEs. Represented industries were retail, technology and services. The variety of backgrounds of the participants was key to getting a wide variety of customers' experiences, which was necessary to understand the different impacts of chatbot accuracy across different industries. The study also selected individuals with experience working in both an agent and a bot capability, so that they could contribute informed comparisons regarding chatbot technology's capabilities versus human agents in actual customer service situations.

Finally, the sample selection took into consideration participants of different levels of technology familiarity, in order to obtain a more balanced view of the matter. Some were deeply conversant with digital technologies, and interacted regularly with chatbots, while others were less so, and came with fewer expectations. This diversity helped with a more detailed understanding of how different user profiles interpret and rate chatbot interactions in terms of accuracy and trust.

3.3. Data Management Plan

All interviews were recorded, transcribed and anonymised to protect anonymity. Names were disguised on transcripts, and identifying details were taken out. The research team had access to encrypted devices where the data were stored securely [1]. First, these procedures helped increase participant trust, and second, the ethical considerations were paramount to making sure participant privacy was maintained. All of the collected data was also cross checked to ensure consistency of the information collected.

To further guarantee data integrity, the transcripts were reviewed independently by several team member. The cross validation process was helpful to minimise bias and to ensure the themes identified represented the experiences of the participants. It was very important in order to give credibility to the study so that the findings reflect the actual experience of the participants.

3.4. Data Analysis Techniques

Interview data were analysed using thematic analysis based on the Braun and Clarke (2006) framework. There were several reasons why thematic analysis was chosen. First, it enables flexible but stringent approach to processing qualitative data, that researchers use to identify patterns and themes, without being bound down by an a priori theoretical frame. This proved useful in understanding the wide array of experiences participants had with chatbot accuracy and customer trust, as the analysis could respond to the individual insights brought by each participant. Thematic analysis was also important because it enabled the in depth exploration of both the positive and negative aspects of chatbot use, so as to provide a balanced view of how it affects customer trust.

Thematic analysis was selected as another reason because it would be appropriate for looking at the experiences of participants that are subjective when discussing topics like trust, satisfaction and frustration. Thematic analysis allows researchers to identify nuances of the meaning participants hold and is therefore an attractive method for the study at hand. Moreover, studies on technology adoption and user experience have made considerable use of thematic analysis, which has a well established methodological foundation that is consistent with the objectives of this research.

The data was familiarised with, coded with significant phrases, themes were identified, and themes were refined to accurately portray participants' experiences. First I read the transcripts multiple times in order to understand the data as a whole. The identified key themes were accuracy in building trust, human fallback and trade-offs between speed and personalisation. Thematic analysis, in its iterative nature, provided opportunities for themes to be refined to more accurately reflect participants' experiences and ultimately offer subjects the voice. This rigorous approach meant that the themes that emerged were consistent across participants but also relevant to the bigger picture of customer service in SMEs.

The coding was done in two phases. Open coding was used in the first phase to identify initial codes as initial codes for the data noted relevant features. Then codes were grouped into categories according to their similarities. In the second phase, these categories were refined with axial coding to determine the relationships between them. This method made sure that all themes discussed were comprehensive and linked in such a way that all experienced was considered holistic. After the themes were defined and named, they were given names to clearly express the meaning of the participants' narratives.

[1] Interview transcripts

4. Results

4.1. Theme 1: Accuracy as the Core of Trust

Consistently, the accuracy of chatbot answers was discussed as a key point for creating trust. Responding accurately and relevantly built trust in the chatbot and the company, participants said. Yet when chatbots gave incorrect, or irrelevant, information participants lost a great deal of confidence. Participants assumed that chatbots would be reliable, especially for simple questions, and inaccuracies were considered to be unacceptable failures which negatively influenced their opinion about the business. The most important reason that participants continued to use chatbot services was accuracy, as consistent and correct answers provided a perception of professionalism and reliability on the part of the company. This finding is consistent with the notion of technological credibility in which accuracy is directly related to the perceived trustworthiness of the service.

What participants said was that even small errors could undermine trust. For instance, one participant recounted a case where a chatbot gave wrong information about item availability, prompting a negative impression of the entire company. For me this emphasizes just how important accuracy is even in seemingly inconsequential exchanges as this determines the customer's overall perception of the brand. It was also mentioned by another participant that they were willing to use a chatbot a second time, but only if the chatbot answered consistently accurately.

4.2. Theme 2: Human Assistance as a Backup

When chatbot limitations were hit, participants insisted that human intervention was necessary. For simple inquiry, chatbots were deemed successful, whereas they often faltered on more complex or nuanced questions. Maintaining trust required the ability to easily transfer the customer to a human agent. Chatbots that could proactively detect when something needed human support and triggered the transfer without any additional customer effort were appreciated by participants. The chatbot showed its readiness to proactively approach this and understood its limitations, and put the customer's needs first. Human backup acted like a safety net; it reduced the frustration the service caused and increased the level of trust by guaranteeing that issues with customers would be resolved, no matter what. This theme is about the necessity of complementing chatbot systems with human support as the second feature instead of by itself resolving all the customer concerns.

One person related an experience with a chatbot that didn't understand what it was asked and would repeatedly give irrelevant responses. However, the participant said the frustration was diminished when they were quickly transferred to a human agent who resolved the issue. The lesson learned from this experience is that seamless human fallback is an absolutely essential component of maintaining trust. One other participant asserted that they would be more likely to make use of chatbots if they knew that in cases of breakdowns, human support would be easily relied on to support any breakdown.

4.3. Theme 3: Frustration with Inaccurate or Generic Responses

A large source of frustration for participants was inaccurate or generic chatbot responses. Participants felt that time was wasted when chatbots gave incorrect answers or gave pre scripted, irrelevant replies. A major pain point was a lack of contextual understanding, which resulted in a negative perception of the chatbot and the business it represented. Participants felt undervalued if the issues that they brought up were not dealt with in the right way, often preferring the human interaction where empathy and contextual understanding was felt to be more reliable. On top of that, chatbots that don't acknowledge their mistakes or provide suitable solutions increase customers' distrust.

One thing that several participants observed was that frustration was always generated by a chatbot because the responses didn't come from personalisation. One participant said they were repeatedly told the same generic thing, and felt like they were not being heard. A second participant told me about a case where a chatbot was unable to comprehend the context of their question prompting a slew of irrelevant responses before they abandoned the conversation. These experiences reinforce why chatbots need to be contextually aware while capable of giving personalized responses so as not to upset the users.

4.4. Theme 4: Expectations vs. Reality

The level of trust was a function of participants' expectations about chatbot capabilities. Participants experienced disappointment when chatbots did not meet their expectations, or increase trust when chatbots exceeded expectations. It was found to be essential to have clear communication about what chatbots can and can not do, thus helping set expectations and avoiding frustration. Some participants said that they had much happier experiences when they knew beforehand about what the chatbot could and couldn't do. Becoming proficient at this reduced the often anticipated disappointment from unrealistic expectations of the capabilities of chatbots. This shows how important it is to have transparency in chatbot design and marketing, so that customers know what chatbot technology is capable of and what it cannot do.

They also mentioned that their initial trust in a chatbot was shaped by how the technology was presented by the company. Other people said when the chatbot was pitched as an advanced solution that could solve complex queries, they expected a lot, and when it didn't work up to expectations, it was very disappointing. However, when participants were told the chatbot was limited, they were more understanding if the chatbot couldn't answer their questions. This leads us to believe that one of the most critical elements in customer trust is that expectations be set as realistic as possible.

4.5. Theme 5: Speed vs. Personalisation

Chatbots were helpful for on-demand, routine tasks you previously had to initiate with a human agent, but in cases of complex or personalized queries, participants liked to talk with a human. Simple questions were seen as a good fit for chatbots and empathetic, context sensitive responses were valued for human agents. The difference prompts the need to find the fine balance between the efficiency of chatbots and the personalisation provided by human service agents. Frequently, participants had mentioned that the speed of chatbots was useful, but at the expense of feeling truly understood. Participants said they, particularly in sensitive or complex situations, valued the human touch, developing a more nuanced understanding of their issues. A key consideration for SMEs was to optimise the customer service while maintaining customer satisfaction, and the trade-off between the speed and personalisation was identified.

Another participant suggested that the advantage of chatbots was that they could give an immediate response, however this also often resulted in the lack of deep responses leaving them unsatisfied. However, human agents were slower to respond but were better able to address their concerns. One of the participants added that the best customer service experience would be having the speed of a chatbot solving simple problems quickly, and the personalization of a human to help with more difficult issues as needed. The feedback to this suggests that SMEs have to be careful of how they deploy chatbots, with customers getting the efficiency they want and the personalisation they require.

4.6. Bar Chart

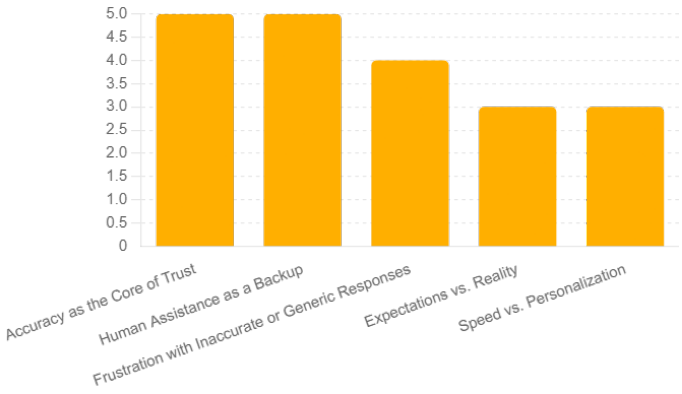

Figure 1. A bar chart which shows the frequency of key themes from interviews, with "Accuracy as the Core of Trust" being the most mentioned and "Speed vs. Personalisation" the least.

For the sake of further illustrating the prominence of these themes, figure 1 was made showing the number of times each theme appeared in the five interviews.

Figure 1 illustrates that Accuracy as the Core of Trust and Human Assistance as a Backup were the most recurring themes, which appear in all five interviews. The themes here highlight the need for chatbot responses that are accurate and for a smooth hand-off to human support when chatbots can't help. These themes are so prominent because chatbots are so efficient but they need to be accurate and defer to human agents when necessary to maintain the trust of the customer.

Less frequently mentioned other themes such as Frustration with Inaccurate or Generic Responses, Expectations vs Reality, and Speed vs Personalisation also help to shape customer perception. Figure 1 shows that Frustration with Inaccurate or Generic Responses was mentioned in four interviews, indicating how negative chatbot experiences can decrease trust. This is particularly important, as this directly determines whether customers will continue to use chatbot services.

Three interviews each had the Expectations vs Reality and Speed vs Personalisation themes, highlighting the need to balance rapid response and personalised service. Figure 1 really makes it clear what issues customers consider when they evaluate chatbot accuracy. This reaffirms that SMEs should tackle such critical elements in the customer service tactics, to achieve overall satisfaction and trust. Businesses can achieve better customer needs and long term trust to their services by focusing on the accuracy of chatbot responses and a smooth transition to human support when needed.

5. Discussion

The results of this study indicate that chatbot accuracy is critical to customer trust in SMEs. Not only does accurate responses add to customer satisfaction, it also helps to give the business a positive image. This is important to note because of the limitations of today's chatbot technology in light of complex or emotionally charged lines of inquiry and the necessity for human fallback support. Automation vs. human support is a critical balance in a good customer service experience. Participants clearly preferred systems that recognised their limitations and effectively managed customer interactions by inserting human agents when needed.

This finding is consistent with previous work highlighting the importance of accuracy and of chatbot qualities that resemble humans (Araujo, 2018; Go and Sundar, 2019). But this study also notes that it is important to manage customer expectations, because unmet expectations can cause frustration and erode the relationship. For SMEs, the key is to have a hybrid approach where chatbot strengths represent just one part of a continuity effort, with human agents in place to address an optimal share of most customer cases in a way that will ensure a minimum level of customer satisfaction and trust. The fact that chatbots can handle simple questions well and humans are good at empathy and problem solving, makes it the perfect model for customer service. Future research could examine which chatbot design features are most effective at generating trust and satisfaction (like conversational tone, empathy cues, and adaptive learning) to understand how to develop such trust and satisfaction building features into chatbot design.

Additionally, it is suggested that chatbot systems continue to need further technological innovations to overcome current limitations. This could improve natural language processing capabilities and chatbots with the added emotional intelligence improvements could better create more real time and empathetic contextually relevant responses. Moreover, future studies can examine whether their customer education can influence the quality of chatbot interactions by indicating customers the functionality and limitations of chatbots, thereby reducing unrealistic expectations, and thereby eventually increasing customers' trust and satisfaction.

6. Conclusion

Finally, to further elaborate, this study presents practical considerations for SMEs when integrating chatbot technology. The main take away from this is that chatbots aren't a complete replacement for human customer service, but rather a complement to it. Using the results of our research to identify areas where chatbots excel (routine and straightforward questions) SMEs can deploy these tools in a way that minimises customer frustration and maximises efficiency. However, it's important to have the ability to escalate cases to human agents when there's more issues to be understood or a more nuanced or emotionally sensitive issue to deal with. This hybrid approach not only satisfies fully the needs of customers but also creates an impression of reliability and responsiveness that can build customer loyalty.

Managing customer's expectation is also another important aspect to consider and that is how transparency plays a role in it. A service that clearly states what a chatbot can and cannot do will be trusted more by the customers. SMEs can partly mitigate the negative impact of unmet expectations of delivering performance that is less than promised by simply setting realistic expectations. It is important to educate customers on what chatbots can do, and offer clear escalation pathways to enable customers to progress their issue to a human. The findings show that transparency is a very powerful tool to build trust, because it aligns customer expectations with the performance of the chatbot system.

Additionally, the study demonstrates the benefits of adding features such as empathy cues and adaptive learning into chatbot design. These empathy cues (expressions of understanding and support) could close the emotional gap that is common in automated interactions. Chatbots can be given such adaptive learning capability which would allow them to get better and better all while learning from past interactions making them more reliable and effective. Since these features generally come with a higher price tag, and SMEs may not have the same need for these features, investing in them could be a sound business move for these organizations since they will differentiate their customer service offering from others' and enhance the customer experience.

Furthermore, there should not be any doubt about the need for on-

going evaluation and feedback loops. SMEs should also regularly assess the chatbot performance and incorporating customer feedback can provide continuous refinements and improvements to chatbot systems. The iterative approach to technology development guarantees that the technology develops on the lines of customer needs and expectations, building on trust and satisfaction.

Finally, chatbots are a great tool, providing efficiency and cost effectiveness but their successful implementation in SMEs needs the perfect balance of technology and human intervention. SMEs can also integrate empathy, transparency and advanced learning capabilities plus a clear human backup strategy in place to create a customer service experience that is not only up to but way beyond customer expectations. Future research should focus on the evolving relationship between chatbot technology and human assistance, seeking to understand what best practises contribute to creating trust, satisfaction, and long term customer loyalty.

References

Gefen, D., & Straub, D. W. (2004). Consumer trust in b2c e-commerce and the importance of social presence: Experiments in e-products and e-services. *Omega*, *32*(6), 407–424. https://doi.org/10.1016/j.omega.2004.01.006

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Oliver, R. L. (2014). *Satisfaction: A behavioral perspective on the consumer*. Routledge.

Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication*, *19*(3), 529–545. https://doi.org/10.1111/jcc4.12066

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior*, *49*, 245–250. https://doi.org/10.1016/j.chb.2015.02.026

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Diederich, S., Brendel, A. B., & Kolbe, L. M. (2019). Designing anthropomorphic enterprise conversational agents. *Business & Information Systems Engineering*, *61*(6), 823–843. https://doi.org/10.1007/s12599-019-00600-8

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304–316. https://doi.org/10.1016/j.chb.2019.01.020

Adam, M., Wessel, M., & Benlian, A. (2020). Ai-based chatbots in customer service and their effects on user compliance. *Journal of Service Management*, *31*(2), 153–187. https://doi.org/10.1108/JOSM-12-2019-0358

Shahid, S., Stirling, D., & Spencer, N. (2020). The impact of chatbots on customer service: A literature review. *Proceedings of the 2020 International Conference on Computing, Electronics & Communications Engineering*, 77–82. https://doi.org/10.1109/iCCECE49321.2020.9231184