SORRY I AM STILL LEARNING – ACTIVE EXPECTATION MANAGEMENT OF CHATBOTS

Martin Haupt¹ and Anna Rozumowski²

¹Technische Hochschule Mittelhessen, Wiesenstraße 14, 35390 Gießen, Germany ²Zürcher Hochschule für Angewandte Wissenschaften, Theaterstrasse 17, 8400 Winterthur, Switzerland

ABSTRACT

Chatbots gained strong popularity in customer service, although users often experienced unsatisfactory interactions – regularly due to highly exaggerated performance expectations. As a viable option, firms might, therefore, think about active expectations management by describing chatbot limitations. However, the question remains whether this strategy has positive or negative impacts on customer satisfaction and re-use intentions. Building on the Expectancy Violation Theory and the CASA-Paradigm, this study plans to evaluate the effects of two different expectation management strategies. As work in progress, we describe the status and plan of this research.

KEYWORDS

Chatbots, Human-Machine-Interaction, Expectation Management, Customer Satisfaction

1. INTRODUCTION

Fueled by recent advancements in technology such as artificial intelligence and machine learning, chatbots have gained popularity (Huang & Rust, 2018). Essentially defined as "text-based agent[s] designed to carry on a conversation" (Schuetzler et al., 2019, p. 253), chatbots deliver information independent of time and space (Bryan, 2019). Therefore, it is not surprising that firms are increasingly implementing chatbots, and the global chatbot market is expected to grow to over 450 million USD in 2027 (The Insight Partners, 2019). Despite this good news, however, users often experience unsatisfactory interaction (Adam et al., 2020). Thus, firms also experienced strong backlashes because of chatbots not performing according to users' expectations. Scholars found that these negative impacts are not only resulting from poor chatbot performance but also that users have exaggerated expectations of the existing chatbots' functionalities (Gentsch, 2019; Klein et al., 2019). To attenuate consumer's expectations to a more realistic chatbot performance level, firms might think about ways to manage user expectations actively. This means that companies could attempt to explain the limited abilities and realistic level of chatbot functions to actively reduce possibly exaggerated expectations. However, management and theory are still unsure, how users would react to this expectation management strategy. In particular, the information about a limited service performance, i.e. a learning chatbot and possibly unsatisfactory answer, might lead either to more understanding or to frustration and chatbot rejections. Moreover, the question is when this information should be provided throughout the conversation. Therefore, we postulate the following research questions:

- Is it possible to actively manage the expectation towards chatbots?
- What influence does the active expectation management have on the interaction with a chatbot?
- Does the timing of expectation management within the conversation influence user reactions?

Building on the Expectancy Violation Theory, this study attempts to demonstrate the effects resulting from different chatbot expectation management approaches. In particular, this study examines how active reduction of expectations towards a chatbot affects consumers' perceptions, attitudes and subsequent behavioral intentions. This research offers several contributions. First, this paper will show that expectations influence consumer reactions regarding chatbot interactions. Second, the study presents and tests "expectation management" as a method to adjust consumer expectations to fit chatbot performance levels. As user expectations were found to be far exceeding general chatbot performance, we analyze how expectation

management could lead to higher chatbot adoption and user satisfaction. Third, to the best of our knowledge, this study is the first to reveal that expectation management effects related to conversations are dependent on the particular position within these conversations. This means an active attempt to adjust a partner's expectations in a conversation before the start of a conversation or during a conversation might lead to different reactions from a communication partner. From a managerial perspective, our study provides guidance for a chatbot design that attenuates high consumer expectations to create more positive user experiences regarding chatbot conversations.

2. CONCEPTUAL FRAMEWORK

Over more than two decades, scholars have investigated human reactions when interacting with machines, such as chatbots or avatars. Nass and colleagues asserted the Computers Are Social Actors Paradigm (CASA), which indicates that people transfer social rules, norms and expectations to interactions with computers (Lee & Nass, 2010; Nass et al., 1996). Several experiments acknowledged this phenomenon and showed that although people were aware of interacting with computers, they ascribed social attributes to them and subsequently treated computers as if they were human (Mou & Xu, 2017). Applying the CASA paradigm to chatbots, implies that a chatbot evokes similar perceptions, conversational norms and outcomes as a chat interaction with a human service agent (Shi et al., 2020). Various studies revealed that a higher level of "humanizing" a chatbot interaction lead to more favorable user reactions, such as satisfaction or re-use intentions (Gnewuch et al., 2017; Qiu & Benbasat, 2014). Likewise, a chatbot's natural conversational abilities enhance user engagement - compared to static information (Zhang et al., 2020). However, this transfer of social norms and a stronger feeling of interacting with a "humanized agent" are causing high expectations towards chatbots (Wirtz et al., 2018). People expect that a chatbot understands their entered information and expect a prompt and suitable answer. Processing of natural language and user text input was found to be a complex task for machines (Bail, 2016). Questions can be posed and answered in manifold ways; this creates difficulties in correctly linking a question to the correct answer (Thomnes, 2020). Furthermore, factors such as ambiguous user input, irony, as well as complex and varying sentence structures create difficulties in the comprehension of user requests (Følstad et al., 2020; Luger & Sellen, 2016). Although current technologies such as artificial intelligence or machine learning are trying to imitate human text understanding, many systems still struggle to interpret information or requests correctly (Gnewuch et al., 2017). Scholars acknowledged that current chatbot systems are quickly reaching their limit and are often failing to handle user requests appropriately (Kvale et al., 2020).

According to the Expectancy Violation Theory (Burgoon, 1993; Burgoon & Hale, 1988), any deviation from a person's currently held expectation will trigger emotions and attempts to explain this behavior. Interpersonal expectancies are hereby described as "enduring pattern of anticipated behavior" (Burgoon, 1993, p. 31) which are affected by different individual, relational and contextual factors. Expectancies help to form cognitive structures to understand and interpret behavior from others; they influence mental information processing and the individuals' behavior. They provide a framework to understand and interpret a partner's behavior in social interactions. This framework is built upon former interactions and experiences, environmental factors and alike. Subsequently, this framework is used to understand a current situation to react appropriately (Ramirez Jr & Zhang, 2007). A violation of the created expectancies may result in positive or negative reactions, depending on the deviation from the personal reference point (Burgoon, 1993). In the field of chatbot interaction, this personal reference point was found to be on a high level as customers compared the chatbot interaction with human service interaction (Gentsch, 2019). These insights support the above-mentioned CASA paradigm. When chatbots do not meet these high expectations, users may experience frustration and dissatisfaction (Luger & Sellen, 2016). To avoid these negative feelings due to a discrepancy between users' expectations and chatbot performance, managers have, in general, two options. Either they close the gap by developing chatbot performance further, or they reduce consumer expectations. Firms are indeed pushing chatbot development, and software firms are continually improving chatbot skills (Schuetzler et al., 2019). However, the latter option has received relatively little attention so far. Initial research showed that users adapt their language when they expect an under-developed (vs. sophisticated) computer system, irrespective of the real performance (Pearson et al., 2006).

3. HYPOTHESES DEVELOPMENT

Reducing user expectations regarding chatbot interactions, reflects the major target of the strategies we labelled as "expectation management". Initial qualitative research proposed that consumers adjust their initial expectations toward chatbot task response (Luger & Sellen, 2016). Furthermore, Burgoon and Hale (1988) described that interaction patterns influence the expectations an individual holds. Therefore, we hypothesize, H1: Expectation management strategies will reduce an individual's expectations toward chatbot performance. Following the Expectancy Violation Theory, the gap between high user expectations and lower service provision creates negative reactions (Burgoon, 1993). Thus, lower expectations would mean that this perceived gap is smaller or even not present, and less negative or even positive reactions occur. Formally, H2: A lower (vs higher) level of expectations should increase (a) user satisfaction and (b) chatbot re-use intentions. Essentially, expectation management might be positioned at two distinct points within the conversation. Either expectations are reduced directly at the beginning of the interaction (ex-ante), or they are reduced when a service failure occurs (ex-post). We expect that both strategies should have differential outcomes. Both strategies should affect users' initial use intention and user's reaction. An expectation reduction before the real conversation (ex-ante) should lead to a reduced feeling of chatbot competence and users might not proceed with the conversation. Thus, H3: Ex-Ante positioning (vs. Ex-Post or none) has a negative effect on chatbot use intention. Additionally, when individuals encounter a service failure during the conversation they might react differently. The ex-ante strategy (vs. ex-post or none) reduces expectations and mentions potential failures before they actually occur. Thus, individuals are expected to feel more prepared and evaluate chatbot conversations with failures more positively. We formulate: H4: Ex-Ante positioning (vs. Ex-Post) has a more positive effect on satisfaction and re-use intention.

4. METHOD

To investigate expectations towards chatbots and the possibilities to guide consumers' expectations, a mixed-methods approach is applied. First, to uncover relevant factors influencing consumer expectancy levels and their attitudinal and behavioral reactions, we will conduct a qualitative study. A focus group discussion with 6 participants will be held. Second, to test and quantify the proposed relationships, we will set up a scenario-based experimental study. We will conduct a between-subjects experiment based on three scenarios (ex-ante vs. ex-post vs. no expectation management). Participants will be randomly assigned to one of the scenarios, where an interaction with a chatbot on a website will be shown. In the first (ex-ante) scenario, the chatbot starts the conversation with additional information about its limited abilities, describing that it is still learning and excuses itself for potential mistakes that it might make. In the second (ex-post) case, the additional information about the chatbot limitations is displayed after the service failure. In the final (none-expectation management) case, the chatbot conversation is shown without any expectation management.

The main body of the chat conversation is kept identical throughout all cases. Three questions are asked; two of which the chatbot answers correctly and one answer involving a service failure. Using established scales from literature, participants are subsequently asked to express their attitudes and behavioral intentions.

5. RESULTS

In this paper, we discuss work in progress. Initial results from a qualitative pre-study (focus group with 6 participants) revealed that people indeed seem to have high expectations towards chatbot's competence, despite a rather bad experience of real chatbot performance. Thus, respondents' expectations seemed to be violated, and this violation was expressed in frustration, anger and decreased trust towards chatbot competence, as well as decreased re-usage intention. Running an experimental study will shed light into customer's use intentions, satisfaction, reactions on the chatbot conversation, and trust towards a chatbot depending on different expectation levels.

REFERENCES

- Adam, M., Wessel, M., & Benlian, A. (2020). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*. https://doi.org/10.1007/s12525-020-00414-7
- Bail, C. A. (2016). Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media. *Proceedings of the National Academy of Sciences of the United States of America*, 113(42), 11823–11828. https://doi.org/10.1073/pnas.1607151113
- Bryan, J. (2019). Bots Gain Importance in Gartner Service Technologies Bullseye. https://www.gartner.com/smarterwithgartner/bots-gain-importance-in-gartner-service-technologies-bullseye/
- Burgoon, J. K. (1993). Interpersonal Expectations, Expectancy Violations, and Emotional Communication. Journal of Language and Social Psychology, 12(1–2), 30–48. https://doi.org/10.1177/0261927X93121003
- Burgoon, J. K., & Hale, J. L. (1988). Nonverbal expectancy violations: Model elaboration and application to immediacy behaviors. *Communication Monographs*, 55(1), 58–79. https://doi.org/10.1080/03637758809376158
- Følstad, A., Araujo, T., Papadopoulos, S., Law, E. L.-C., Granmo, O.-C., Luger, E., & Brandtzaeg, P. B. (Eds.). (2020). *Chatbot Research and Design*. Springer International Publishing. https://doi.org/10.1007/978-3-030-39540-7
- Gentsch, P. (2019). Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business -- Konzepte und Best Practices (2., überar). Springer Fachmedien. https://doi.org/10.1007/978-3-658-25376-9
- Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards Designing Cooperative and Social Conversational Agents for Customer Service. Proceedings of the International Conference on Information Systems, 38.
- Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. Journal of Service Research, 21(2), 155–172. https://doi.org/10.1177/1094670517752459
- Klein, R., Schulz, M., Reinkemeier, F., Dill, F., Schäfer, A., & Hoffmann, P. (2019). *Chatbots im E-Commerce: Desillusion oder großes Potenzial* (D. P. Elaboratum (Ed.)). https://www.chatbot-studie.de/
- Kvale, K., Sell, O. A., Hodnebrog, S., & Følstad, A. (2020). Improving Conversations: Lessons Learnt from Manual Analysis of Chatbot Dialogues. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, & P. B. Brandtzaeg (Eds.), *Chatbot Research and Design* (Vol. 11970, pp. 187–200). Springer International Publishing. https://doi.org/10.1007/978-3-030-39540-7 13
- Lee, J.-E., & Nass, C. (2010). Trust in Computers: The Computers-Are-Social-Actors (CASA) Paradigm and Trustworthiness Perception in Human-Computer Communication. *Trust and Technology in a Ubiquitous Modern Environment: Theoretical and Methodological Perspectives*, 1–15. https://doi.org/10.4018/978-1-61520-901-9.ch001
- Luger, E., & Sellen, A. (2016). "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.*"
- Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. Computers in Human Behavior, 72, 432–440. https://doi.org/10.1016/j.chb.2017.02.067
- Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? International Journal of Human-Computer Studies, 45(6), 669–678. https://doi.org/10.1006/ijhc.1996.0073
- Pearson, J., Hu, J., Branigan, H. P., Pickering, M. J., & Nass, C. I. (2006). Adaptive language behavior in HCI. In G. Olson & R. Jeffries (Eds.), *Proceedings of the SIGCHI conference on Human Factors in computing systems -CHI '06* (p. 1177). ACM Press. https://doi.org/10.1145/1124772.1124948
- Qiu, L., & Benbasat, I. (2014). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, 25(4), 145–182. https://doi.org/10.2753/MIS0742-1222250405
- Ramirez Jr, A., & Zhang, S. (2007). When Online Meets Offline: The Effect of Modality Switching on Relational Communication. *Communication Monographs*, 74(3), 287–310. https://doi.org/10.1080/03637750701543493
- Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2019). The effect of conversational agent skill on user behavior during deception. *Computers in Human Behavior*, 97, 250–259. https://doi.org/10.1016/j.chb.2019.03.033
- Shi, W., Wang, X., Oh, Y. J., Zhang, J., Sahay, S., & Yu, Z. (2020). Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–13. https://doi.org/10.1145/3313831.3376843
- The Insight Partners. (2019). Chatbot market revenue worldwide from 2018 to 2027 (in million U.S. dollars). https://www.statista.com/statistics/1007392/worldwide-chatbot-market-size/
- Thomnes, J. (2020). Chatbots & Co. Künstliche Intelligenz steckt noch in den Kinderschuhen (Horizont Verlag (Ed.)). https://www.horizont.net/tech/nachrichten/chatbots--co-kuenstliche-intelligenz-steckt-noch-in-den-kinderschuhen-185748
- Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: service robots in the frontline. *Journal of Service Management*, 29(5), 907–931. https://doi.org/10.1108/JOSM-04-2018-0119
- Zhang, J., Oh, Y. J., Lange, P., Yu, Z., & Fukuoka, Y. (2020). Artificial Intelligence Chatbot Behavior Change Model for Designing Artificial Intelligence Chatbots to Promote Physical Activity and a Healthy Diet: Viewpoint. J Med Internet Res, 22(9), e22845. https://doi.org/10.2196/22845