Chatbots in Stressful Customer Service Situations – The Impact of Stress on the Intention to Use Conversational Chatbots

Tobias Heußler
Wiesbaden Business School

Cite as:
Heußler Tobias (2021), Chatbots in Stressful Customer Service Situations – The Impact of Stress on the Intention to Use Conversational Chatbots. *Proceedings of the European Marketing Academy*, 50th, (93611)

Paper from the 50th Annual EMAC Conference, Madrid, May 25-28, 2021

Chatbots in Stressful Customer Service Situations – The Impact of Stress on the Intention to Use Conversational Chatbots

Abstract

Chatbots enhance the efficiency of several customer touchpoints in terms of the customer experience and associated costs. However, what happens in stressful situations? Do customers still use chatbots in such cases? Research on the usage of chatbots has focused on technological aspects. By contrast, this paper contributes to a research stream that analyzes the acceptance of chatbots. This research reflects stress as an influencing factor in the technology acceptance model (TAM) and examines the moderating role of stress on usefulness and their relationship on the intention to use chatbots. In an experimental design a first group of subjects talked to a call center agent. A second group was pretended to talk to a chatbot. Actually, this group also talked to an agent, whose voice was manipulated by a voice changer. The empirical results demonstrate that the impact of stress needs to be considered when employing chatbots for daily interactions with customers.

Keywords: Chatbots, acute stress, behavioral intention to use

Track: Digital Marketing

1. Introduction

Many companies today are competing in markets where the channels of customer interactions are changing rapidly (Huang and Rust, 2017). Recently, chatbots as interactive touchpoints have gained increasing attention in customer communication on websites and have become even more appealing in mobile apps or microprograms (Araujo, 2018; Longoni & Cian, 2020).

Chatbots or conversational agents are defined as computer programs that communicate with users using natural language, respond automatically to language or text in a human-like manner, and execute specific commands (Thomaz, Salge, Karahanna and Hulland, 2020). Implemented perfectly, these conversations will be similar to those users have with their friends and family.

Marketing managers appreciate chatbots for two major reasons: (1) Chatbots help companies maintain governance over direct customer interfaces. As such, they can enhance the customer experience, collect customer data, and win back margins paid to retailers and platform suppliers (Thomaz et al., 2020; Verhagen, Van Nes, Feldberg and van Dolen., 2014; Rese, Ganster, and Baier, 2020). (2) Chatbots are more cost efficient than service personnel (Lester, Mott, and Branting, 2005). Due to the growing levels of automated communications, experts expect even higher savings potential in the future (Thomaz et al., 2020). However, managers fear they will require a high capital expenditure and result in low customer acceptance (Verhagen et al., 2014).

Although knowledge on customer reactions to chatbots is indispensable to achieving higher customer satisfaction and effective savings, this aspect has rarely been addressed in the marketing literature (Brandtzaeg and Følstad, 2017). Therefore, this study enhances our understanding of situations where consumers evaluate the interactions with chatbots as valuable and acceptable. We assess the intention of chatbot usage in the context of stressful situations and, in turn, identify conditions when engagement with service staff may be more appropriate than chatbots. The phrasing of our hypotheses is based on the technology acceptance model (TAM) (Davis, Bagozzi, and Warshaw, 1989). Finally, the results of the ANOVA provide the implications of a "successful" chatbot implementation and ways to improve the efficiency of customer service.

2. Conceptual Model

Marketing literature on chatbots and their willingness to use has, to date, concentrated on cognitive influencing factors and on the technical performance of chatbots (Chung, Ko, Joung, and Kim, 2018; Thomaz et al., 2020). Most authors refer to the TAM2 by Venkatesh and Davis (2000) as an appropriate framework. TAM2 is an extension of the TAM by Davis, Bagozzi, and Warshaw (1989) and theorizes about the usage of new technologies as chatbots. Among other variables, perceived usefulness and ease of use determine the behavioral intention to use (BI). Perceived usefulness can be defined as "the extent to which a person believes that using a system will enhance his or her […] performance" (Venkatesh and Davis, 2000: 187) and could be affected by external variables. Finally, the actual usage depends on the BI (Venkatesh and Davis, 2000). Over the last, 20 years, several authors have confirmed the model, even in online shopping and service environments (Legris, Ingham, and Collerette, 2003). By contrast, this research draws on interactive elements in communication (Mimoun Poncin, and Garnier, 2017; Wang and Benbasat, 2008). To arrive at a more accurate understanding of chatbot acceptance, we need to broaden the range of investigated determinants. A few authors have analyzed affective determinants (Chung et al., 2018; Zarouali, Van den Broeck, Walrave, and Poels, 2018). Thus far, only Zarouali et al. (2018) confirmed affective influencing factors such as pleasure, arousal, and dominance (PAD-dimensions) in determining the likelihood to use chatbots.

However, the role of stress in the context of chatbots has been neglected so far (Rese, Ganster, and Baier, 2020). The above-mentioned study did not even control for stress. This lack of attention to stressful situations is surprising since, for decades, a large body of consumer research on stress has shown that individuals tend to change their behavior in their interactions with companies and their service staff (Mathur, Moschis, and Lee, 2006; Singh and Duque, 2012). Moschis (2007) revised the relationship between stress and consumer behavior. He distinguished between acute and chronic stress. His study focused on acute stress because acute stress occurred in situations in which consumers were confronted with new technologies. Acute stress is defined as “discrete, observable events which are thought to be threatening because they represent change” (Wheaton, 1990: 210). Previous studies have confirmed that bias judgment due to acute stress impacts consumer spending (Durante and Laran, 2016), price sensitivity (Maier and Wilken, 2014), purchase abandonment (Albrecht, Hattula, and Lehmann, 2017), and service encounters (Singh and Duque, 2012). According to powerful

psychological stress theories, acute stress disrupts the consumers' equilibrium. Buying situations may create instability among inner forces (Moschis 2007). If one's equilibrium is disrupted, then there are strategies to restore the equilibrium and to cope with stressful situations (De Jong Gierveld and Dykstra, 1993): problem solving (rational decision making), avoidance (cognitive or behavioral), tension-reduction behaviors (e.g., exercising), social support (from family and friends), information seeking (from the professional community), and situation redefinition (viewing the situation differently and diminishing its perceived severity). In the context of buying decisions assisted by chatbots, we assume that avoidance is a relevant strategy to maintain the equilibrium. In turn, the intention to use chatbots will decline.

Combining the considerations of Moschis (2007) on stress with the TAM2, we propose that stress interacts the behavioral intention to use chatbots (BI). A negative stress state reduces the intention to use chatbots in contrast to well-known and reliable service channels (e.g. personal assistants, call center agents). Hence, we add to the literature by proposing the following hypothesis:

H: Acute stress impacts the intention to use a chatbots more negatively than the intention to use a personal service channels (here: call center agent).

3. Empirical Study and Results

3.1 Sample design and experimental scenario

Our sample comprised 220 students from a German university, including 116 women and 104 men. The average participant age was 23.8 years. The students were invited to an IT-laboratory of the university. The stimuli consisted of scenarios describing the booking of a train trip. Since students are familiar with decisions in the mobility and tourism sector, research has confirmed ticket booking as a high-involvement scenario, which has been tested in several experimental designs (Meloni, Sanjusta, Sottilea and Cherchib, 2013; Bamberg, 2000).

A short introduction to the basic scenarios was given to the participants by a scientific instructor: To buy (refund) a train ticket for your friend, you visit a website. In order to book (refund) the train ticket, you need to go through the booking engine. However, the booking engine does not work properly, so you are asked to talk to a chatbot (or a call center agent)." Participants got specific booking requirements: origin/destination (O/D), train times, train classes, and a reason for the refund of the ticket (illness of the friend).

Finally, participants started a real conversation via phone. A first group talked to a human call center agent. A second group was pretend to talk to a chatbot. In both cases, participants

talked to a human call center agent. However, in the case of the chatbot we adjusted the voice via a voice changer. The voice sounded machine-made.

We did not pay incentives to participants as the decisions in our experiments were not evaluated in terms of right or wrong and, thus, had no basis for performance-oriented payments (Homburg, Hoyer, and Koschate, 2005).

3.2 Manipulation, measures and results

Our study used a 2x2 between-subjects design (Hair, Black, Babin und Anderson, 2018), crossing the service channel (chatbot and call canter agent) and the acute stress state (neutral, stressful). The dependent measure was the intention to use the channel again. To control confounding variables, we assigned subjects to the experimental groups randomly (Homburg, Hoyer, and Koschate, 2005).

The manipulations for acute stress began with a review of the existing literature, which offers little guidance (Maier and Wilken, 2014; Durante and Laran, 2016). Moschis (2007) identified two major sources of stress in purchasing situations. Stressful consumer decisions are expected to be a result of either an unexpected set of products to choose from or limited time for decision making (Moschis, 2007). Hence, we changed the stress level by giving the subjects a strict time limitation of four minutes. To pretest the level of stress, we conducted different scenarios in terms of time restrictions. Additionally, we increased the number of possible travel options in terms of transport modes and prices.

As proposed by Herr, Page, Pfeiffer, and Davis (2012), we did not include the manipulation checks in the experiment as this would have biased the behavior. We conducted a pretest measuring the skin response rate and using a self-report in the laboratory conditions (Wang and Minor, 2008; Maier and Wilken, 2014). Two electrodes measured the skin conductivity caused by sweat. Stress leads to higher degree of perspiration, and, in turn, increases skin conductivity. According to Boucsein (2012) measuring the skin response rate is appropriate as long as the measurement last throughout the complete pretest to level out the skin response rate for every single participant. As a result, participants in the stressful scenario have more skin conductions than in a neutral condition. The difference was significant ($t=10.42$; $p<.001$).

In the experiment we added a self-report of stress developed by Mehrabian and Russel (1974). This has been used by marketing scholars to assess stressful responses to interpersonal aspects of shopping (Maier and Wilken, 2014). Respondents indicated the level of stress on a scale of 1 (not stressful at all) to 10 (very stressful) on the stress/relaxation continuum (Mehrabian and Russel, 1974). The manipulation in the experimental groups perceiving

stressful and neutral states was confirmed by the scale (Mean Stress=6.35; Mean Neutral=3.96). A pairwise t-test between the stressful and neutral scenarios was large in the intended direction (t=12.83; p<.001).

After finishing the call, subjects were confronted with the evaluation of the behavioral intention to use (BI) either the chatbots or the agent . Consistent with the elaborated scales of the TAM (Venkatesch and Davis, 2000), we measured BI by using a four-item scale from 1 ("fully disagree") to 7 ("fully agree"). The mid-point of the scale was described as "neither." All scales had acceptable reliability (Cronbach's Alpha=.86). In addition, exploratory factor analysis of the scale returned a single factor solution.

We hypothesized that the behavioral intention to use chatbots or call center agents differs and is interacted by the stress level. The results strongly support this hypothesis. To test the significance of the differences, a 2x2 ANOVA using BI as the dependent measure revealed the desired interaction effect (F (1; 219) = 23.747; p<.001). This effect is presented in Figure 1. The results indicated that a high stress level increase the importance of the offered sales channel. In the context of a neutral stress level the BI to use either a chatbot or a human agent was of minor importance. This is surprising, as the result of the conversation (ticket refunded) has been the same. From a rational standpoint it does not matter whether participants talked to the chatbot or the human agent.

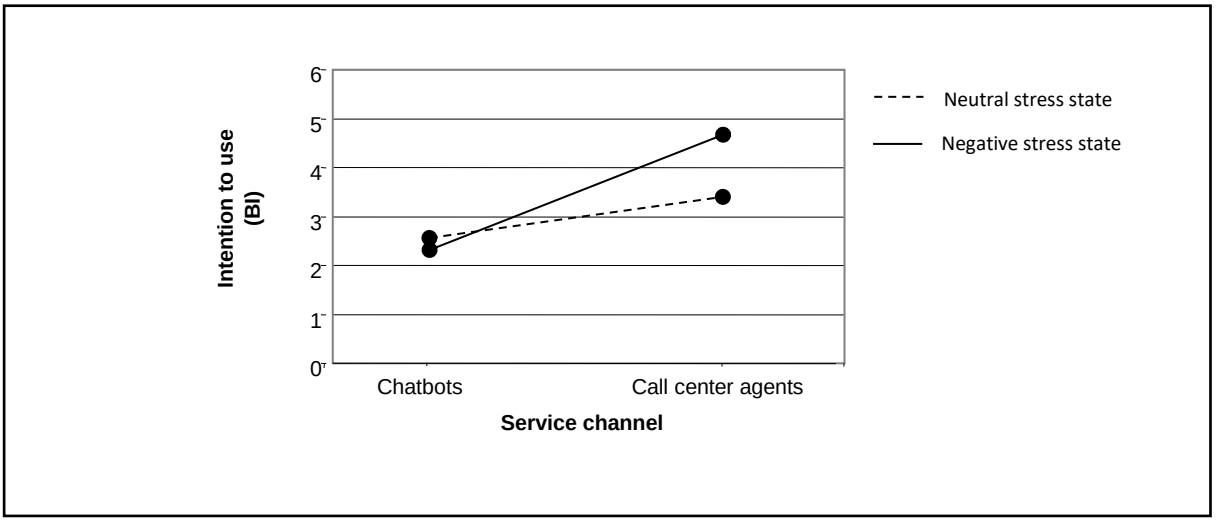

Fig. 1: Means of behavioral intention (BI) by experimental condition

4. Covid-19 pandemic, Limitations and Outlook

Although most managers believe that chatbots will have a positive influence on the company's operational financial performance, they fear low customer acceptance (Thomaz et al., 2020). Thus, our study extends previous research in the area of chatbots by investigating how

stress in buying situations influences and moderates the intention to use chatbots. This accompanies cognition considerations on perceived usefulness or ease of use within the TAM. In fact, this study generally supports the increased attention in the marketing literature on stress, as cognitive theories have been shown to have limitations in explaining consumer behavior (Moschis 2007).

As the experiment has been conducted prior to the COVID-19 pandemic, we need to reconsider our results in light of the COVID-19 pandemic. Based on theoretical reasoning we thereby assume that the results should generally hold stable. However, one could speculate that people get more used to digital sales and communication channels due to the restrictions in the lock-down phases. Consequently, the reluctance to use chatbots may decline even in stressful situations. Indeed, we started to replicate the experiment in October 2020. However, unfortunately, the number of participants is small (36). Hence, there are no significant changes to be seen.

In addition to COVID-19 pandemic, there are other approaches for future research. Chatbot development currently relies on improving technological aspects (e.g., using artificial intelligence to improve conversational quality) (Thomaz et al. 2020), while stress analysis, as part of general acceptance, has been neglected thus far. Generally, our study shows that marketing managers need to pay more attention to the role of stress in all aspects of communication, particularly, when developing chatbots as service or sales channels. However, our study focused only on neutral and stressful states. Further research should extend the spectrum of stress and investigate whether there are other reasons for stress and how these could be addressed through counter measures. In addition, we did not measure the stress caused by the use of the chatbot itself.

Thus, a major question for managers is how to reduce the stress level in buying or service situations? And more concretely in the context of our experimental setting, in how far does human assistance and the intervention in conversations help reduce stress, particularly, in emotional conversations? Future research may touch on this idea of emotional elements within conversations (O'Neill and Lambert, 2001). As our findings show, the total replacement of human agents, especially in respect to emotionally stressful situations, does not seem promising yet. Nevertheless, subjects do not deprecate the use of chatbots in neutral stress states.

More generally, it seems promising to analysis the deeper reasons for the decline of the intention to use chatbots in stressful situations. There are a lot of speculative reasons. Subjects may trust in humans more than in bots. Or, subjects simply feel left alone with technology,

resulting in a perceived lack of appreciation by the service provider. Having found the major reasons, it would be easier for managers to find the right balance between an appreciative and automated communication.

There are three more limiting aspects that need to be addressed to enhance the explanatory power of our research. First, our model controls exclusively for stress. Hence, the model could be extended to other aspects related to technology (e.g., privacy, or speed of conversation) or behavior (e.g., trust). Second, we tested a voice-based, not a messenger chatbot. Our model could be used with messenger chatbots to see if there are substantial differences between the two types of chatbots (Elkins and Derrick, 2013). Third, we need to consider whether chatbots and their technological state limit their potential and spread. Even in the near future, we expect that their general acceptance will increase (Thomaz et al., 2020). Thus, the analysis of our findings over time will be promising.

As with any methodology, there are limitations associated with our experimental research. The results refer to a limited data set. Additionally, the use of a booking scenario is not complex enough to generalize the findings. Consequently, future research should explore the relationships within the TAM among other consumer groups and buying situations.

References

Albrecht, C.-M., Hattula, S., & Lehmann, D. R. (2017). The relationship between consumer shopping stress and purchase abandonment in task-oriented and recreation–oriented consumers. *Journal of the Academy of Marketing Science*, 45 (5), 720–740.

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.

Bamberg, S. (2000). The promotion of new behavior by forming an implementation intention: Results of a field experiment in the domain of travel mode choice, *Journal of Applied Social Psychology*, 30 (9), 1903–1922.

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. In I. Kompatsiaris, I. Cave, A. Satsiou, G. Carle, A. Passani, E. Kontopoulos, D. Sotiris, & D. McMillan (eds.), *Internet Science* (pp. 377–392). Cham, Springer.

Boucsein, W. (2012). *Electrodermal activity* New York: Springer.

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2018). Chatbot e–service and customer satisfaction regarding luxury brands. *Journal of Business Research*, 117 (9), 587-595.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35 (8), 982–1003.

De Jong Gierveld, J., & Dykstra, P. A. (1993). Life transitions and the network of personal relationships: Theoretical and methodological issues, *Advances in Personal Relationships*, 4, 195-227.

Durante, K. M., & Laran, J. (2016). The effect of stress on consumer saving and spending. *Journal of Marketing Research*, 53 (5), 814–828.

Elkins, A. C., & Derrick, D. C. (2013). The sound of trust: voice as a measurement of trust during interactions with embodied conversational agents. *Group decision and negotiation*, 22 (5), 897–913.

Hair, J. F., Black, W. C., Babin, B., & Anderson, R. E. (2018). Multivariate Data Analysis, 8th edn, Englewood Cliffs, Pearson.

Herr, P. M., Page, C. M., Pfeiffer, B. E., & Davis, D. F. (2012). Affective in-fluences on evaluative processing. *Journal of Consumer Research*, 38, 833–845.

Homburg, C., Hoyer, W. D., & Koschate, N. (2005). Customer reactions to price increases: Do customer satisfaction and the perceived motive fairness matter? *Journal of the Academy of Marketing Science*, 33 (1), 36–49.

Huang, M.-H., & Rust, R. T. (2017). Technology–driven service strategy. *Journal of the Academy of Marketing Science*, 45 (6), 906–924.

Legris, P., Ingham, J., & Collerette, P. (2003). Why do people use information technology? A critical review of the TAM. *Information & Management*, 40 (3), 191–204.

Lester, J., Branting, K., & Mott, B. (2005). Conversational agents. In M. P. Singh (ed.), *The Practical Handbook of Internet Computing* (pp. 220–240). Baton Rouge, Chapman & Hill.

Longoni, C., & Cian, L. (2020). Artificial Intelligence in Utilitarian vs. Hedonic Contexts: The "Word-of-Machine" Effect. *Journal of Marketing*, November (2), 1–18.

Maier, E., & Wilken, R. (2014). The impact of stress on consumers' willingness to pay. *Psychology & Marketing*, 31 (9), 774–785.

Mathur, A., Moschis, G. P., & Lee, E. (2006). Consumer stress-handling strategies: theory and research findings. *Journal of Consumer Behaviour: An International Research Review*, 5 (3), 93–203.

Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. Boston, MA: MIT Press.

Meloni, I., Sanjusta, B., Sottilea, E., & Cherchib, E. (2013). Propensity for voluntary travel behavior changes: An experimental analysis. *Procedia–social and Behavioral Sciences*, 87, 31–43.

Mimoun, M. S. B., Poncin, I., & Garnier, M. (2017). Animated conversational agents and e–consumer productivity: The roles of agents and individual characteristics. *Information & Management*, 54 (5), 545–559.

Moschis, G. P. (2007). Stress and consumer behavior. *Journal of the Academy of Marketing Science*, 35 (3), 430–444.

O'Neill, R. M., & Lambert, D. R. (2001). The emotional side of price. *Psychology & Marketing*, 18 (3), 217–237.

Rese, A., Ganster, L., & Baier, D. (2020). Chatbots in retailers' customer communication: How to measure their acceptance?. *Journal of Retailing and Consumer Services*, 56, 102–176.

Singh, S., & Duque, L. C. (2012). Moderating role of stress in evaluating negative services: Encounters with the police. *Journal of Service Research*, 15 (2), 231–241.

Thomaz, F., Salge, C., Karahanna, E., & Hulland, J. (2020). Learning from the Dark Web: leveraging conversational agents in the era of hyper-privacy to enhance marketing. *Journal of the Academy of Marketing Science*, 48, 1–21.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46 (2), 186–204.

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.

Wang, Y. J., & Minor, M. S. (2008). Validity, reliability, and applicability of psychophysiological techniques in marketing research. *Psychology & Marketing*, 25, 197–232.

Wang, W., & Benbasat, I. (2008). Attributions of trust in decision support technologies: A study of recommendation agents for e–commerce. *Journal of Management Information Systems*, 24 (4), 249–273.

Wheaton, B. (1990). Life transitions, role histories, and mental health. *American Sociological Review*, 55, 209–223.

Zarouali, B., Van den Broeck, E., Walrave, M., & Poels, K. (2018). Predicting consumer responses to a chatbot on Facebook. *Cyberpsychology, Behavior, and Social Networking*, 21 (8), 491–497.