Custom-Made Chatbots.

About the Potential of active User-Chatbot-Customization for the User Experience Design

Rebecca Wald[1,2][0000-0003-2086-903X], Evelien Heijselaar[2], and Tibor Bosse[2][0000-0003-4233-0406]

[1] University of Amsterdam, Amsterdam School of Communication Research, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands
[2] Radboud University, Behavioural Science Institute, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands

r.wald@uva.nl

Abstract. Chatbots have become a worldwide phenomenon and find application in many different domains. However, their inability to reliably succeed in social contexts (e.g., on Social Network Sites, in therapy) significantly influences the user's experience. To improve user experience developers and scholars have mostly concentrated on strategies that originate from the developers' side. Less attention has been paid to the active involvement of the user and to potential customization options. Several psychological mechanisms suggest active customization to facilitate user engagement with the technology and potentially cushion conversational issues that still widely exist in today's chatbot interactions. Reviewing existing literature and related empirical findings on this topic, we advocate the position that active User-Chatbot-Customization has great potential for user experience design and the improvement of user engagement in particular. Going forward, certain theoretical, empirical and methodological aspects discussed here must be overcome in order to further exploit this potential. Directions for future research are outlined.

Keywords: chatbots, user experience design, customization, user engagement

1   Introduction

Chatbots are promised to be of great economic potential and societal value [1, 2, 3]. Subsequently, they have received growing attention by businesses and scholars likewise and are expected to play a major role in future digital communication. Chatbots can be defined as virtual conversational agents that utilize Natural Language Processing to engage with their user in a text-based format "the same way humans communicate with each other" [4, p. 1]. Chatbots that are specifically designed for the social context

(e.g. on Social Network Sites, in therapy) and for potential long-term relationship formation aim to create an interpersonal connection with their user, similar to a companion or friend [5]. For those chatbots it is crucial to be able to socially engage with their user as if they represent an actual human conversation partner. To the time of this writing, research, however, finds the opposite [4, 6]. Users often complain about the chatbots' poor use of human-like communication style, about their scriptedness, and about their quality of advice giving [7, 8]. In therapeutic contexts specifically, users criticize chatbots' inability to express adequate empathy and authenticity [9]. This, consequently, weakens their user experience and makes future encounters not only more complicated (e.g., stronger resistance to reveal personal information, poorer persistence in case of communicational errors), but also less likely in general [e.g., 4, 8, 10, 11, 12, 13].

To again enhance user experience and the likelihood of user retention, developers typically utilize strategies such as altering the chatbot's language to compensate for mistakes and errors (e.g., politeness, formal apologies), integrating anthropomorphic cues (e.g., human-like avatar name and look), supporting the conversational agent with a (real) human agent [14], or providing more user-orientated content [15]. The latter is commonly known under the term of *personalization* that is based on user input and is implemented in current chatbot development suites such as Dialogflow or Azure Bot Service.

What those strategies all have in common is that they represent solutions that originate from the developers' side only mostly without any active user involvement. Often times, the result is one chatbot for all users. Yet, a generic solution fails to account for individual differences that empirical studies have identified. Firstly, users choose to engage with a chatbot for multiple reasons. Some want to experience fun and entertainment, others strive to satisfy social and relational needs (e.g. emotional support, reducing anxiety), and again others simply want to follow their curiosity about the technology [5, 16]. Secondly, researchers have discovered individual differences in the degree to which users are willing and able to perceive the chatbot as a mindful social interaction partner [17]. Lumping users together must therefore be done with care, since the unique set of user preferences, expectations, and preconditions loosens the ground for the common strategies mentioned above to work as intended. For instance, a user who feels easily threatened by a human-like chatbot does not benefit much from additional anthropomorphic cues; if anything, those cues can backfire and let the chatbot fall into the uncanny valley of the mind (see [18] for a detailed explanation of this effect). Similarly, a user who is looking for a serious conversation with a chatbot gets more easily frustrated if the chatbot responds in a rather humorous and sarcastic way. Hence, an authentic and meaningful chatbot experience becomes increasingly unlikely.

2 Aim and Objective

The limitations of generic user-experience strategies highlight the importance of novel features that are able to strengthen the interpersonal connection between the user and the chatbot. In that, the growing chatbot market is still missing a *personal chatbot*

[19] that is able to generate intimate knowledge about its user and is customizable by the user herself *for* herself. The objective of this paper is to introduce an alternative approach to the traditional one-chatbot-fits-all solution, namely *active User-Chatbot-Customization*. Here, *active* refers to the involvement of the user, whereas *customization* refers to "the degree to which [the] technology […] or service can be created, selected, or changed [by the user] to comply with user preferences" [20, p. 1549]. Consider the following scenario: A user decides to engage with a chatbot on a Social Network Site out of a specific precondition, namely feeling lonely. This precondition translates into a specific need, for example the need for social interaction. The need for social interaction further elicits the expectation from the chatbot to make oneself not feeling so lonely anymore. A generic chatbot would appear to and interact with this user in the same way it would do with every other user. It would have a certain type of chatbot avatar photo, a specific persona, and language style. In contrast, a chatbot that is customizable by the user along her preferences (e.g., preference for a male conversation partner with mutual interests and humorous language skills) (1) would form a better fit, (2) would allow the user to feel more personally engaged, since the user is actively involved in the setting of those characteristics, and (3) would eventually increase the chance for the user to satisfy her expectations.

To substantiate this alternative approach, we review related literature that can provide first empirical insights and theoretical reasoning for the study of customization in chatbots. Furthermore, we aim to outline directions for future research to follow up on this, since exploiting the hidden potential of active User-Chatbot-Customization requires more in-depth and large-scale empirical testing.

3   Related Work

Up to the point of writing, customization has received most empirical attention in the context of gaming research. Particularly, with the focus on user-avatar identity and self-representation [21]. Figure 1 illustrates an example of a customization interface through which a player can choose characteristics of a game avatar.

Fig. 1. An example of the Sims avatar customization interface [22].

Respective research has found that customizing the appearance and abilities of a game avatar has led players to report more fun and user engagement [23]. Scholars and game designers trace this back to the so-called immersion effect that originates in the coalescences between the real and virtual world [24]. Since a game avatar is supposed to visually represent the real-world player in the game's virtual world, the immersion effect is an expression of the degree to which the player believes to be physically and mentally present in the game. In contrast, for a chatbot interaction the immersion effect is reversed: the virtual character (i.e., the chatbot) is supposed to immerse into the real world, allowing the chatbot user to have a natural and realistic conversation. First indications for the immersion effect to work bidirectionally were discovered by Dolgov et al. [25]. The researchers found game players, who were able to customize their game avatar, to be more likely to identify with their in-game player in a real-world situation after completing the game. Although this forms fruitful space for empirical investigations of immersion effects in chatbot interactions, empirical efforts to test this remain scarce.

To our knowledge, only one other study has previously investigated the effects of a customized chatbot. Xiao et al. [26] utilized a 2×2 balanced between-subject design through which they found that creating the illusion of being able to customize a chatbot's characteristics and appearance made the user rate the agent "as more lik[e]able, more trustworthy and more useful" [26, p. 1299]. Yet, further empirical efforts are strongly needed to replicate and extent these findings.

Regarding the theoretical understanding of customization and potentially underlying psychological mechanisms, literature from multiple disciplines provides several starting points. The Computers Are Social Actors (CASA) paradigm [27] – prominent within the field of Human-Computer Interaction – suggests that humans generally respond to computers in a social manner and apply human-specific concepts even though they are aware of the non-human origin of their counterpart. If users are likely to treat a chatbot similarly to how they treat other human conversation partners, it seems logical that users should respond positively to human cues and subsequently attribute a human

mind to the chatbot. Though, empirical findings do not fully allow such generalizations [16]. Thus, it is necessary to consult additional literature that focuses more strongly on individual users and their needs.

For this, the discipline of communication science provides the Uses and Gratifications (U&G) theory [28] which claims that users "actively" choose a medium to fulfil their individual needs. When combining this theoretical proposal with the CASA paradigm, it is to assume that a chatbot is likely to be seen and treated as a human conversation partner, however every user does so independently and based on personal motives. This corresponds to the user behavior previously illustrated in our scenario.

Lastly, the fields of psychology and marketing research introduce the so-called IKEA effect. This effect explains customers' greater willingness to consume, engage with, and benefit from products through self-creation [29]. Following this understanding, the process of customizing a chatbot could potentially strengthen interpersonal connection and closeness between user and chatbot, foster user engagement, and increase the benefit of the interaction, simply because the user is able to co-create the chatbot. Hence, self-creation makes it more likely for the user to get from the chatbot what she needs.

Despite these first explanations for positive customization effects, a comprehensive theoretical understanding is still lacking. One of the key tasks for future research is therefore to develop such a framework. The following section proposes ideas for respective empirical investigation.

4 Future Research

Empirical research of chatbot customization seems fruitful in a wide variety of contexts. Among those especially the ones that locate the locus of control already on the user side (e.g. customer support) or allow the development of long-term relationships (e.g. personal assistant or counsellor) are well suited [30]. In the health care domain, for example, customizable chatbots could present a supporting tool for mental health counselling and therapeutic treatment [31]. Patients would be able to have more control over their conversation partner and would therefore be more likely to feel more invested in the conversations and able to trust the given advice. The implementation of User-Chatbot-Customization could thus form a promising pathway to substitute a human counsellor in the future.

The biggest challenge for empirical investigations, however, is that we find ourselves in a scenario of infinite regress. While there does not yet exist a chatbot software that is customizable by users without programming skills [19], a plausible reason for this non-existence is the scarcity of empirical evidence for potential social and economic benefits. Instead of waiting until a respective application appears on the market to be tested more thoroughly, we want to encourage researchers to aid technological development through the use of creative methodologies that allow pioneering findings in this research niche.

A possible way to qualitatively study chatbot customization is via visual demos [32] that subsequently are discussed in focus groups or think-aloud studies. Showcasing a sample interaction with a customizable chatbot would allow participants to envision themselves as a user of a customizable chatbot without the need for the technology to exist already. An exploratory approach is especially suitable here as it enables researchers to analyze little known social phenomena, like the effect of self-creation in chatbot design, in short time and at low cost [33].

To further quantify new discoveries, researchers can design vignette studies that describe a hypothetical chatbot encounter and potential options for customization. Respective measurements for use intention, enjoyment, and usefulness can be implemented as evaluation metrics to test conceptual models derived from theory. Nevertheless, it must be noted that conclusions drawn from demos or vignettes only partially touch upon the actual concept of chatbot customization. Researchers must keep in mind that analyzing users' reactions and evaluations of a hypothetical chatbot interaction is not the same as doing so based on the actual chat experience.

To go beyond a hypothetical scenario, future research can also utilize a Wizard of Oz approach [34]. Here, participants are made to believe that they are talking to a computer, while actually a human – the wizard – guides the conversation. This way, customization features (e.g. setting the look, language style, personal interests of the chatbot) can be manually incorporated in an already existing chatbot application (e.g. Mitsuku, [35]) through persona-scripts prepared a-priori. A follow-up survey can finally assess users' chatbot experience. Yet, here too, scientific compromises are made. Particularly, the number of degrees of freedom increases to the point that outcomes become hardly comparable due to (1) differences in the conversational flow between participants and (2) the additional variance created through the experimenter as the human in the loop. This approach can eventually weaken empirical findings and must be considered when interpreting the results.

5   Conclusion

As a valuable contribution to the growing body of knowledge about chatbot user experience design, this position paper outlines the disadvantages of a one-chatbot-fits-all-users approach, while providing suggestions for the implementation of active User-Chatbot-Customization in empirical chatbot research. Before the potential of this alternative approach can be further exploited, future scholars need to (1) develop a more comprehensive theoretical understanding of individual differences between chatbot users, to (2) identify underlying (psychological) mechanisms suggested by various scientific disciplines, and to (3) apply and develop robust methodologies to reliably test assumptions and draw meaningful conclusions. For this, designers and scientists must work in symbiosis and advance scientific research and technological development in parallel.

References

1. Følstad, A., Brandtzaeg, P. B., Goodwin, M., Granmo, O.-C.: Chatbot for loyalty: From technology push to human touch. In: Proceedings of the international workshop on chatbot research CONVERSATIONS 2017, Thessaloniki, Greece, https://conversations2017.files.wordpress.com/2017/11/conversations_2017_paper_8_folstad.pdf, last accessed 2020/09/11 (2017).
2. Kepuska, V., Bohouta, G.: Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: Proceedigns of IEEE 8th Annual Computing and Communication Workshop and Conference, CCWC, pp. 99-103 (2018).
3. Luger, E., Sellen, A.: "Like having a really bad pa": The gulf between user expectation and experience of conversational agents. In: Proceedings of CHI'16, May 07 - 12, 2016, San Jose, CA, USA, pp. 5286–5297. https://doi.org/10.1145/2858036.2858288 (2016).
4. Skjuve, M., Brandzaeg, P. B.: Measuring user experience in chatbots: An approach to interpersonal communication competence. In: Skjuve, M., Brandtzaeg, P. B. (eds.) Internet Science, LNCS, pp. 113-120. Springer, Heidelberg (2019).
5. Shum, H., He, X., Li, D.: From Eliza to XiaoIce: challenges and opportunities with social chatbots. Frontiers of Information Technology & Electronic Engineering, 19(1), pp. 10-26. http://orcid.org/0000-0002-4684- 911X (2018).
6. Chaves, A. P., Gerosa, M.: How should my chatbot interact? A survey on human-chatbot interaction design. Manuscript submitted to ACM. (2018).
7. Jain, M., Kumar, P. Kota, R., Patel, S. N.: Evaluating and Informing the Design of Chatbots. In: Proceedings of the 2018 Designing Interactive Systems Conference June 2018, pp. 895–906. https://doi.org/10.1145/3196709.3196735 (2018).
8. Følstad, A., Brandtzaeg, P. B.: Users' experiences with chatbots: Findings from a questionnaire study. Quality and User Experience, 5(1), pp. 1–14. https://doi.org/10.1007/s41233-020-00033-2 (2020).
9. Bell, S., Sarkar, A., Wood, C.: Perceptions of chatbots in therapy. In: Proceedings of the conference on Human Factors in Computing Systems, 1–6. https://doi.org/10.1145/3290607.3313072 (2019).
10. Chatbots Magazine, https://chatbotsmagazine.com/chatbot-report-2019-global-trends-and-analysis-a487afec05b, last accessed 2020/09/11 (2019).
11. Kulms, P., Kopp, S.: More human-likeness, more trust?: The effect of anthropomorphism on self-reported and behavioral trust in continued and interdependent human-agent cooperation. In: Proceedings of MuC '19, September 8–11, 2019, Hamburg, Germany, pp. 31–42. https://doi.org/10.1145/3340764.3340793 (2019).
12. Piccolo L.S.G., Mensio M., Alani H.: Chasing the Chatbots. In: Bodrunova S. et al. (eds) Internet Science. INSCI 2018. Lecture Notes in Computer Science, vol 11551. Springer, Cham. http://doi-org-443.webvpn.fjmu.edu.cn/10.1007/978-3-030-17705-8_14 (2019).
13. Zamora, J.: I'm Sorry, Dave, i'm afraid i can't do that: Chatbot perception and expectations. In: Proceedings of the 5th International Conference on Human Agent Interaction HAI '17, Bielefeld, Germany, pp. 253–260. https://doi.org/10.1145/3125739.3125766 (2017).
14. Grudin, J., Jacques, R.: Chatbots, Humbots, and the Quest for Artificial General Intelligence. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3290605.3300439 (2019).
15. Fan, H., Poole, M. S.: What is personalization? Perspectives on the design and implementation of personalization in information systems. Journal of Organizational Computing and Electronic Commerce, 16(3&4), pp. 179-202. https://doi.org/10.1080/10919392.2006.9681199 (2006).

16. Brandtzaeg, P. B., Følstad, A.: Why people use chatbots. In Kompatsiaris et al. (Eds.). Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, pp. 377–392. https://doi.org/10.1007/978-3-319-70284-1_30 (2017).

17. Lee, S., Lee, N., Sah, Y. J.: Perceiving a Mind in a Chatbot: Effect of Mind Perception and Social Cues on Co-presence, Closeness, and Intention to Use. International Journal of Human-Computer Interaction, 36(10), pp. 930 – 940. https://doi.org/10.1080/10447318.2019.1699748 (2019).

18. Stein, J.-P., Ohler, P.: Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. Cognition, 160. https://doi.org/10.1016/j.cognition.2016.12.010 (2017).

19. Daniel, F., Zaccaria, V., Matera, M., Dell'orto, A.: Toward truly personal chatbots. On the development of custom conversational assistants. In: Proceedings of ACM/IEEE. https://doi.org/10.1145/3195555.3195563 (2018).

20. Teng, C.: Customization, immersion satisfaction, and online gamer loyalty. Computers in Human Behavior, 26(6), pp. 1547–1554. https://doi.org/10.1016/j.chb.2010.05.029 (2010).

21. McArthur, V.: The UX of Avatar Customization. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems May 2017, pp. 5029–5033. https://doi.org/10.1145/3025453.3026020 (2017).

22. Carl's Guide: The Sims 4 How To's. URL: https://www.carls-sims-4-guide.com/tutorials/how-to/make-a-sim.php, last accessed: 2020/11/06. (2020).

23. Turkay, S.: User experiences with avatar customization in Second Life and Lord of the Rings online. In: PRoceedigns of Teachers College Technology Conference, 2012, 78 (2012).

24. Goethe, O.: Immersion in Games and Gamification. In: Goethe, O. (ed.) Gamification Mindset, pp. 107-117 Springer, Switzerland. https://doi.org/10.1007/978-3-030-11078-9 (2019).

25. Dolgov, I., Graves, W. J., Nearents, M. R., Schwark, J. D., Brooks Volkman, C.: Effects of cooperative gaming and avatar customization on subsequent spontaneous helping behavior. Computers in Human Behavior, 33, pp. 49–55. https://doi.org/10.1016/j.chb.2013.12.028 (2014).

26. Xiao, J., Stasko, J., Catrambone, R.: The role of choice and customization on users' interaction with embodied conversational agents: Effects on perception and performance. In: Proceedings of the conference on Human Factors in Computing, pp. 1293–1302. https://doi.org/10.1145/1240624.1240820 (2007).

27. Nass, C., Steuer, J., Tauber, E. R.: Computer are social actors. In: Proceedigns oft he Conference on Human Factors in Computing Systems. CH194-4/94 Boston, Massachusetts USA, pp. 72–78 (1994).

28. McQuail, D.: The rise of media of mass communication. In: McQuail, D. (Ed.), Mass communication theory: An introduction, pp. 1–29. London: Sage (1994).

29. Mochon, D., Norton, M. I., Ariely, D.: Bolstering and restoring feelings of competence via the IKEA effect. International Journal of Research in Marketing, 29(4), pp. 363-369. https://doi.org/10.1016/j.ijresmar.2012.05.001 (2012).

30. Følstad, A., Skjuve, M., Brandtzaeg, P. B.: Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design. Proceedings of the international workshop on chatbot research CONVERSATIONS 2018, St. Petersburg, Russia. (2018).

31. Cameron, G., Cameron, D., Megaw, G., Bond, R., Mulvenna, M., O'Neill, S., Armour, C., McTear, M.: Towards a chatbot for digital counselling. In: Proceedings of British HCI 2017 –Digital. http://dx.doi.org/10.14236/ewic/HCI2017.24 (2017).

32. Candello, H., Pinhanez, C., Pichiliani, M. C., Alberio Guerra, M., Gatti de Bayser, M.,: Having an Animated Coffee with a Group of Chatbots from the 19th Century. In: Proceedigns

of CHI EA '18: Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, April 2018, Paper No.: D206, pp. 1–4. https://doi.org/10.1145/3170427.3186519 (2018).

33. Acocella, I.: The focus groups in social research: advantages and disadvantages. Qual Quant, 2012(46), pp. 1125–1136. https://doi.org/10.1007/s11135-011-9600-4 (2012).

34. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of Oz studies - why and how. Knowledge-Based Systems, 6(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N (1993).

35. Pandorabots: Meet *Mitsuku*. https://www.pandorabots.com/mitsuku/, last accessed: 2020/11/06. (2019).