Chatbot Design Features to Increase Productivity

Lukas Ramming[1][0000-0003-3415-1029], Daniel Konadl[1][0000-0002-3703-8024] and Susanne Leist[1][0000-0002-2376-1027]

[1] University of Regensburg, Germany
{Lukas.Ramming, Daniel.Konadl, Susanne.Leist}@ur.de

Abstract. In recent years, chatbots have become a growing presence in our everyday lives. Companies have identified various potential use-cases posing opportunities for reducing costs and providing services through automating processes with the help of chatbots. However, although an increasing number of chatbots are developed, user expectations often cannot be met, leading to frequent discontinuation of the bots. Research suggests that for users, one of the main reasons to use a chatbot is to help them increase their productivity. The literature base so far provides little prescriptive knowledge guiding implementation of chatbots specifically for use-cases where productivity is the main purpose. This short paper is the first step within a Design Science Research project to close this gap. We conducted a systematic literature review and gathered chatbot Design Features that were covered in respective publications. In the further course of our project, we intend to implement a prototype chatbot and formulate Design Principles to provide prescriptive knowledge for designing productivity oriented chatbots.

Keywords: Chatbots, Design Science Research, Rigor Cycle, Systematic Literature Review, Design Features, Productivity.

1 Introduction

Chatbots have become a well-known presence in the digital realm, and we encounter them in various contexts. Companies use chatbots for automating certain customer service tasks, providing users instant responses to their requests [1]. To clarify terms, with chatbots we refer to software programs that mimic human-to-human dialogues by receiving users' utterances (e.g., via text messages or via button clicks), processing the utterances and subsequently presenting a corresponding response to the user [1]. According to [2], one of the most frequent motivational factors for using chatbots is increased productivity, with chatbots providing task assistance and access to information in a timely manner. However, in spite of technological advancements, chatbots still fail to meet users' expectations, leading to dissatisfaction among users and discontinuation of chatbot projects [3, 4]. While the amount of research on chatbot design has increased tremendously in recent years [5], there is only little research on chatbot design that is aimed specifically at increasing productivity. One of the few studies about the topic is by [6] who measured productivity of users who completed different tasks with the chatbot assisting them. The productivity measure used by the authors is based on [7] and will also be used to operationalize productivity for our research project. Accordingly,

productivity is the ratio of output to input. In the context of chatbots, the output can be defined as the task or the number of tasks completed and the input as the amount of time needed for the completion [6]. In our project, we focus specifically on the potentials of chatbot design to increase productivity. We do not consider further types of solutions (e.g., self-service portals) within this research endeavour. Moreover, we do not yet focus on a specific application domain or specific types of tasks as we intended to gather a broad range of DFs based on a comprehensive analysis of literature. We formulate the following research question: *Which chatbot Design Features should be implemented in a context where increased productivity is a desired outcome?* In the next section, we describe the research procedure. We then present the results from our systematic literature review and conclude with an outlook on next steps within the Design Science Research (DSR) project.

2 Research Procedure

To investigate the research question, we follow the DSR approach laid out by [8]. While our overall DSR project will cover all three cycles of this approach, in this paper, we cover the Rigor Cycle as a first step. Thereby, we ground our project in the existing knowledge base by reviewing literature regarding the general topic of chatbot productivity. Following this paper, in the second step, we advance to the Relevance Cycle. With the help of practitioners in the field, we firstly want to identify a specific application and task domain and secondly, extend and prioritize the list of Design Features (DF) that we identified throughout this short paper for the specified domain. In the third step, we enter the Design Cycle and implement a productivity increasing chatbot for the specified domain and with identified DFs to derive Design Principles (DP) [9], providing prescriptive knowledge [10] to chatbot designers in a setting where high productivity is desired. Further details of the next steps following this short paper are described in section 4. To move forward in the Rigor Cycle, we conducted a systematic literature review following the methodological approaches of [11] and [12]. We focused our research on contributions that discuss chatbot design and provide recommendations or present empirical evidence of effects of DFs on goal variables related to productivity. The brackets accompanying the searched databases contain details on the respective keyword-based search queries as follows: (Hits per database | Investigated hits per search query | Papers identified as potentially relevant). Thereby, the following six databases and search engines were applied: *ScienceDirect* (4.358 | 250 | 22), *SpringerLink* (3.579 | 250 | 15), *ACM Digital Library* (1,040 | 250 | 24), *AIS Electronic Library* (18,067 | 250 | 21), *IEEE Xplore* (228 | 228 | 17) and *Google Scholar* (21,700 | 250 | 58). We used the following keywords to find potential hits for our literature review: *(("(conversational OR digital OR virtual) AND (agent OR assistant)") OR "chatbot") AND ("efficient" OR "efficiency" OR "success" OR "successful" OR "effective" OR "effectiveness" OR "productive" OR "productivity")*. This initial literature research resulted in a total of *157 papers*. Regarding our research question, we only kept articles relating to productivity-related DFs of chatbots, based on their title, the provided keywords and their abstract (*remaining: 86*). Moreover, duplicates found within the

keyword-based searches were eliminated (remaining: 78). We also applied language (English), thematic focus (covering chatbot design) and quality (peer-reviewed) as inclusion criteria (*remaining: 64*). In a second step, we also conducted a forward as well as a backward search that resulted in another 8 articles. Therefore, a total of *72 papers* were selected for in-depth analysis in connection with our literature review.

3   Results: Chatbot Design Features and Categories

The systematic literature review uncovered a broad range of chatbot DFs potentially influencing productivity measures. To delimit terms, a DF describes a specific chatbot design element that provides a certain functionality. A DP on the other hand, is a concept in DSR and specifies prescriptive design knowledge according to a specific schema [9]. Studies that directly investigated the effect of certain DFs on productivity as defined in the introductory chapter could not be identified. Therefore, we selected DFs that respective literature argues to be beneficial to the chatbot interaction in more general terms - either quantitatively by positively affecting certain target variables or qualitatively, identified as favourable through user feedback. A list of DFs extracted from the identified papers was created. To reduce overlaps, similar or identical DFs were merged independently by two of the authors resulting in a consolidated list of DFs. During our analysis, we found that the features mainly fall into only a few categories. To define these categories, we followed the inductive category development approach as proposed by [13]. Thereby, two researchers involved within this DSR project independently derived the categories described in Tables 1 - 3. DFs mapped to category 1 are summarized in Table 1. These are DFs that augment the textual input paradigm of a chatbot by implementing features that are well-know from applications based on a graphical user interface (GUI) like, for example, quick responses (e.g., buttons). We argue that these elements, through speeding up the conversation [14, 15], can increase productivity by decreasing the required input (e.g., the amount of time, by reducing user's typing efforts). In the following, the sources provided in Tables 1 – 3 are to be seen as exemplary sources that cover respective DFs. For a full list of the 72 papers that the DF identification is based on, see [16].

Table 1. Design Features augmenting the textual input paradigm with GUI elements

DF #	Description	Source
DF1	Presenting quick responses (e.g., buttons, carousels) for faster navigation and less typing effort by the users.	[14, 15, 17]
DF2	Providing conversational context through separate GUI elements (e.g., buttons, text fields) that allow manipulation of the context.	[17, 18]
DF3	Providing dynamic quick responses taking into account the conversational context and showing the user possible next steps.	[6, 15, 19]
DF4	Using rich media elements like, for example, images, GIFs and videos in the chat widget to convey information in a visual way.	[15, 17, 20]
DF5	Allow users to edit previous messages by clicking and altering them, thus possibly updating the current intent and the conversational context.	[19, 20]
DF6	Decrease typing efforts by implementing a text prediction feature (auto-completion) providing word and intent suggestions.	[6, 17]
DF7	Provide a persistent menu with chatbot capabilities that is always available.	[15, 17]

The second category of DFs we could identify support the conversational process. They are intended to decrease conversational breakdowns that research shows can lead to lower user satisfaction and users abandoning the specific task [4, 17]. Although a direct effect of this category of DFs on productivity may not be as obvious as for DFs from the first category and secondary effects are to be expected, we assume they have an overall positive effect on productivity by reducing conversational breakdowns and therefore increasing the number of completed tasks. Taking DF11 as an example, a misunderstanding between user and chatbot can lead to repeated questions and answers without reaching a conclusion. DF11 can help to reformulate the input by providing specific words leading to the misunderstanding and therefore increase the chance of a successful task completion leading to increased productivity. On the other hand, user may find calls for reformulation disturbing and they could also slow down the conversation leading to lower productivity measures. Overall, although the DF can increase productivity, the exact effect depends on the specific context and implementation of the DF. In the further course of our research, we will address this line of questioning.

Table 2. Design Features supporting the conversational process

DF #	Description	Source
DF8	Providing information about the conversational context in chatbot responses to bring the user and the chatbot on a common ground of understanding. (e.g., tell the current intent and entities the bot has identified).	[17, 18]
DF9	Identify different signs indicating that the user might abandon the interaction soon (e.g., repeated reformulations) and provide alternatives (e.g., human handover).	[19, 21]
DF10	In case of misunderstandings, provide the user the chatbot capabilities and the intents that have the highest probability based on the user's input so far.	[4, 17, 21, 22]
DF11	In case of misunderstandings, reflect the user message back to the user and highlight signal words (e.g., words that could not be understood or that led to a certain classification of input) and ask for reformulation of the request.	[4, 21, 22]
DF12	When starting the conversation, introduce the chatbot as artificial and provide the user the chatbot capabilities to help manage user expectations.	[17, 21]

There is a broad literature base covering anthropomorphic chatbot design and social cues in chatbot conversations. Anthropomorphic chatbot design can elicit social responses in users, positively affecting dependent variables like service satisfaction [23, 24]. This could also have a positive effect on productivity if users are more inclined to use a chatbot. However, a humanized chatbot interacting in a social way (e.g., by engaging in small talk) [25], may also distract users from completing their task. Moreover, there has been research suggesting that DFs like buttons, that make a chatbot interface more like a common GUI application, negatively influence the social presence of the chatbot [14], which could also have an effect on productivity measures. Therefore, we also want to consider the effect of this category of DFs (Table 3) on productivity.

Table 3. Design Features personifying the chatbot

DF #	Description	Source
DF13	Using social cues (e.g., emojis, response delays, name etc.) to achieve a certain level of humanness.	[20, 26]
DF14	Choosing a consistent conversation style that supports the context of task orientation.	[17, 25]

4   Next Steps and Limitations

As outlined in the second chapter, the next step of our DSR project, will cover the Relevance Cycle [8]. In this cycle we focus on two goals. Firstly, we want to narrow the scope - so far chatbot productivity in general – to a specific application domain and a specified set of tasks to be solved. Secondly, we will prioritize and extend the list of DFs we identified based on the requirements of the specified application domain. To achieve both goals we will conduct interviews with practitioners in the field. Firstly, to ensure that our DSR project contributes to solving an important and relevant problem and secondly, to gather DF that have not been covered in literature yet and rank DFs in respect to their expected effects on productivity, based on practitioners' experience from actual chatbot projects. Concerning the Design Cycle, we will instantiate a productivity oriented chatbot for the specified domain, implementing the most promising DFs. Conducting experiments with a between-subject design (cf. [27]), we want to extract the effect of specific features on productivity. Eventually, through an iterative process of going through the three cycles, we want to formulate DPs [9] to provide prescriptive knowledge to designers implementing a chatbot in a specific domain where productivity is desired. These DPs can serve as the bases for future research about increasing chatbot productivity in other domains. The following limitations must be mentioned: Our definition of productivity introduced in chapter 1 includes input and output as determinants of productivity. However, there are further measures that can influence productivity and could be included in a more comprehensive productivity definition (e.g., cognitive load). Future research can identify which measure are connected to chatbot productivity and examine their effect. Another direction for future research is to study interaction effects of DFs from the different categories (see Tables 1 – 3) for a deeper understanding of their potential impact on productivity. The contribution of this short paper is the systematic literature review with the resulting overview of chatbot DFs having the potential to increase productivity.

References

1. Stucki, T., D'Onofrio, S., Portmann, E.: Chatbot – Der digitale Helfer im Unternehmen: Praxisbeispiele der Schweizerischen Post. HMD (2018)
2. Brandtzaeg, P.B., Følstad, A.: Why People Use Chatbots. INSCI (2017)
3. Janssen, A., Grützner, L., Breitner, M.: Why do Chatbots fail? A Critical Success Factors Analysis. ICIS (2021)
4. Ashktorab, Z., Jain, M., Liao, Q.V., Weisz, J.D.: Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns. CHI (2019)
5. Diederich, S., Brendel, A.B., Morana, S., and Kolbe, L.: On the design of and interaction with conversational agents: An organizing and assessing review of human-computer interaction research. JAIS 23(1), (2022)
6. Schmidhuber, J., Schlogl, S., Ploder, C.: Cognitive Load and Productivity Implications in Human-Chatbot Interaction. ICHMS (2021)
7. Fried, H.O., Schmidt, S.S., Lovell, C.A.K.: The Measurement of Productive Efficiency. Techniques and Applications. Oxford University Press (1993)

8. Hevner, A.: A three cycle view of design science research. Scandinavian Journal of Information Systems 19(2), (2007)

9. Gregor, S., Kruse, L., Seidel, S.: Research Perspectives: The Anatomy of a Design Principle. JAIS 21(6), (2020)

10. Gregor, S., Hevner, A.: Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37(2), (2013)

11. Webster, J., Watson, R.: Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26(2), (2002)

12. vom Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R., and Cleven, A.: Reconstructing the giant: On the importance of rigour in documenting the literature search process. ECIS (2009)

13. Mayring, P.: Einführung in die qualititative Sozialforschung. Eine Anleitung zu qualitativem Denken. Beltz Verlag (2002)

14. Diederich, S., Brendel, A.B., Lichtenberg, S., Kolbe, L.M.: Design for fast request fulfillment or natural interaction? Insights from an experiment with a conversational agent. ECIS (2019)

15. Valério, F.A.M., Guimarães, T.G., Prates, R.O., Candello, H.: Chatbots Explain Themselves: Designers' Strategies for Conveying Chatbot Features to Users. JIS 9(3), (2018)

16. Full Reference List Available on GitHub:. https://github.com/chatbot-research/desrist22

17. Jain, M., Kumar, P., Kota, R., Patel, S.N.: Evaluating and Informing the Design of Chatbots. DIS (2018)

18. Jain, M., Kota, R., Kumar, P., Patel, S.N.: Convey: Exploring the Use of a Context View for Chatbots. CHI (2018)

19. Li, C.-H., Yeh, S.-F., Chang, T.-J., Tsai, M.-H., Chen, K., Chang, Y.-J.: A Conversation Analysis of Non-Progress and Coping Strategies with a Banking Task-Oriented Chatbot. CHI (2020)

20. Klopfenstein, L.C., Delpriori, S., Malatini, S., Bogliolo, A.: The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms. DIS (2017)

21. Toxtli, C., Monroy-Hernández, A., Cranshaw, J.: Understanding Chatbot-mediated Task Management. CHI (2018)

22. Benner, D., Elshan, E., Schöbel, S., Janson, A.: What do you mean? A Review on Recovery Strategies to Overcome Conversational Breakdowns of Conversational Agents. ICIS (2021)

23. Diederich, S., Brendel, A.B., Kolbe, L.M.: Designing Anthropomorphic Enterprise Conversational Agents. BISE 62(3), (2020)

24. Gnewuch, U., Morana, S., Adam, M., and Maedche, A.: Faster is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. ECIS (2018)

25. Chattaraman, V., Kwon, W.-S., Gilbert, J.E., Ross, K.: Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. CHB 90, (2019)

26. Feine, J., Gnewuch, U., Morana, S., Maedche, A.: A Taxonomy of Social Cues for Conversational Agents. IJHC 132, (2019)

27. Boudreau, M.-C., Gefen, D., Straub, D.W.: Validation in Information Systems Research: A State-of-the-Art Assessment. MIS Quarterly 25(1), (2001)