

Customer Service Chatbots: A Qualitative Interview Study into Customers' Communication Journey

Margot J. van der Goot¹[0000-0001-6904-6515], Laura Hafkamp², and Zoë Dankfort²

¹ Amsterdam School of Communication Research (ASCoR), Nieuwe Achtergracht 166, 1018 WV Amsterdam, the Netherlands
m.j.vandergoot@uva.nl

² Ruigrok Netpanel, Silodam 1A, 1013 AL Amsterdam, the Netherlands

Abstract. The current qualitative interview study describes the communication journey of customers who wish to contact companies, and their evaluation of chatbot communication within this journey. Interviews were conducted with a sample (N=24) that was varied in terms of gender, age, educational level and household composition. Experiences with nine customer service chatbots were included. The analysis focuses on three stages in the journey: first, customers' prior expectations when contacting a company; second, their experiences during chatbot conversations, and third, their final conclusions about under which conditions customer service chatbots should be implemented, and the consequences of chatbot communication for customers' company perceptions. Implications for research and practice are discussed.

Keywords: AI, Chatbots, Company perceptions, Customer service, Qualitative interview study, User experience

1 Introduction

Developments in AI fundamentally alter how companies communicate with their customers [9]. Particularly in customer service, chatbots are increasingly implemented [7, 11, 12]. Customers who need information or who want to complain can type their questions in a dialogue screen (often looking like a chat interface), and receive answers in natural language. The essential characteristic of this type of communication is that, although the answers are automatically generated, the conversation is made to resemble a dialogue between humans [7]. Since this is a novel way of interacting with a company, the question arises how customers experience these conversations, and how this communication affects their company perceptions. The current qualitative interview study aims to shed light on these issues by describing the communication journey of customers who wish to contact companies, and their evaluation of chatbot communication within this journey. The analysis focuses on three stages in the journey: first, customers' prior expectations when contacting a company; second, their experiences during chatbot conversations, and third, their final conclusions about under which conditions customer service chatbots should be implemented, and the consequences of chatbot communication for customers' company perceptions.

The current analysis is most closely related to recent qualitative work that also focused on users' perceptions of customer service chatbots [5, 6, 13]. As Følstad and Skjuve noted, thus far, research providing insights in user experiences and motivations regarding customer service chatbots has been severely limited. The present analysis extends the previous qualitative studies in at least two ways. First, previous studies have already shown that the prime motivation for the use of customer service chatbots is to get their (simple) customer queries answered in a fast and convenient matter [6, 13]. The present study takes this as given, and provides further context to this motivation by showing how communication with customer service chatbots is embedded in customers' communication journey with a company. Second, compared to the interview study in which customers were interviewed after their chatbot use of two specific companies [6], the present study encompasses more variation. The current study uses nine chatbots, of both profit and non-profit organizations. The chatbots are also varied in terms of humanlike characteristics. Moreover, whereas the strength of Følstad and Skjuve's study was that it approached people who actually used the two chatbots, the present study takes the approach of inviting a wider variety of people. The current sample includes both people who use customer service chatbots as well as people who are rather unwilling to use them.

By doing so, this study provides further insight into real-life experiences customers have with existing chatbots, thus providing necessary context to experimental work that investigates effects of -particularly- humanlike cues on persuasion outcomes and company perceptions [2, 9, 10]. The study also provides guidance to companies that are faced with the challenge of implementing chatbot technology for their customer service. Although much has been written about how to design and scale chatbots, it remains important to look at chatbot communication through customers' eyes.

2 Background

This section showcases what previous academic research has already shown about the three stages in the communication journey that we focus on in the current study.

2.1 Customers' Prior Expectations of Communication with a Company

Previous qualitative work has shown that the primary reason for customers to enter a conversation with a customer service chatbot is to get their (simple) customer queries answered in a fast and convenient manner [3, 5, 6, 13]. This implies that the key to successful chatbot communication in customer service is that the queries are answered in a correct and fast way. This makes sense, but in real life this motivation is embedded in the broader journey of communicating with a company. Therefore, the current qualitative interview study sets out to describe what the expectations are that customers have when contacting a company, and what the position is of chatbot communication within the journey.

In addition to this main motivation for using customer service chatbots, research has shown that the mood people are in may determine their experiences during the chatbot conversation. Particularly, Hadi found that customers who entered the chat in an angry

emotional state reported lower satisfaction with the chat when it was humanized compared to when it was not. This in contrast to non-angry customers for whom humanization could actually enhance their customer satisfaction [9]. This suggests that we need to know more about the states that customers are in when they enter a conversation with a customer service chatbot. Therefore, the current study asks:

RQ1: What do customers expect from communication with a company, how do they use customer service chatbots within this communication journey, and what moods are they in when they start a chatbot conversation?

2.2 Customers' Experiences during Chatbot Conversations

Once customers decide to use a chatbot, the question arises what the conversation features are that optimize their user experience. The design aspect that received most attention in design education as well as in academic research is the humanization of chatbots [6, 7]. Humanizing chatbots can be done by adding anthropomorphic cues such as a name, persona, and by using a conversational language style with dialogical cues [8].

Quite some experimental studies have manipulated specific anthropomorphic cues in chatbots to assess the effects of such cues [2, 8, 9]. For instance, Araujo investigated whether anthropomorphic design cues influenced perceptions about the chatbot as well as company-related outcomes [2]. He found that humanlike language or a name were sufficient to increase the perception of the agent as being humanlike. This experiment also showed that the usage of humanlike cues had a significant influence on the emotional connection customers felt with the company. Thus the study found some initial evidence that chatbots with humanlike cues can have a positive effect on relationship building. However, Hadi found that the outcomes of humanizing chatbots depend on customer characteristics, but also on the specific service context. Specifically she found that humanization of the chatbot indeed improved customer satisfaction, but not when customers were angry [9].

It is relevant to place this humanization in the broader context of the overall user experience during a chatbot conversation. In the aforementioned qualitative study, Følstad and Skjuve found that customers experienced humanization and language style as less important than whether they received help with their enquiries [6]. Therefore, qualitative research is called for to further delve into the question of what role humanization plays in the user experience. The present study does this by asking the following question:

RQ2: What are the most important characteristics of customers' experiences with customer service chatbots? And, more specifically, how do they experience the humanization of these chatbots and their language style?

2.3 Overall Evaluations of Customer Service Chatbots and Company Perceptions

The aforementioned qualitative studies provided some preliminary insights in how customers overall evaluate the implementation of chatbots for customer service. Følstad and Skjuve concluded that customers have quite realistic expectations of what customer service can and can not do [6]. However, these studies did not specifically look into relations between perceptions of the chatbot's performances and company perceptions.

Some previously mentioned experimental studies did look into the effects of (specifically) humanizing chatbots on company perceptions. Araujo found initial evidence that humanlike cues can have a positive impact on the emotional relation that customers feel with a company. However, such effects were not found for attitudes towards the company [2]. Also, as said, Hadi found that these effects of the humanization of chatbots on evaluations of the company depend on the emotional state customers are in [9]. The current study builds on this previous qualitative and experimental work by asking:

RQ3: What are customers' final conclusions about under which conditions customer service chatbots can be implemented, and what are the consequences of chatbot communication for customers' company perceptions?

3 Method

3.1 Interviews

The qualitative interview study was approved by the Ethics Review Board of the first author's university, and the interviews were conducted by ISO-certified research agency Ruigrok Netpanel. The idea for this study came from the first author, and the other two authors are researchers/interviewers employed by Ruigrok Netpanel. Each interview was conducted by one of these two interviewers, whereas the other two researchers were present in the adjacent observation room and discussed the ongoing interview to check whether there were any additional questions that needed to be asked. The interviews took place on three days in November 2019, on two locations in the Netherlands (Amsterdam and Amersfoort). All interviews were conducted in Dutch, lasted one hour, and participants signed a consent form prior to the interview. All interviews were video recorded, and transcribed verbatim.

Interview Guide. The interview guide was developed by the three researchers, and was slightly adjusted between the three days. The interview guide consisted of an introduction and three topics. The introduction was mainly an introduction of the agency, the collaboration with the university, and contained some guidelines regarding how the interview would work, but did not yet contain information regarding the specific interview topic. The first topic focused on communication with companies. The interviewer asked whether the interviewee had contacted companies before, in what type of situations, for what type of questions, through what types of communication channels, and with what expectations. We also wanted to know whether interviewees spontaneously mentioned chatbots; therefore the interviewer did not mention chatbots, virtual agents etc. in any way here. The responses to these questions are used to answer RQ1.

After this first topic, the interviewee was asked to interact with one of the chatbots that we preselected. The second topic pertained to their experiences during this specific

chatbot conversation. First, the questions were open-ended and did not yet pertain to specific characteristics of the conversation. The interviewer asked “what did you just do?”, “what happened here”, “how did it go”, and “to what extent do you have the feeling that you are helped?”. This enabled us to answer the first part of RQ2.

Subsequently, the interviewer aimed to gain insight in the experiences with the anthropomorphic cues and the language style, which is the second -more specific- part of RQ2. The interviewer first probed into what the interviewee thought he/she was communicating with, by asking “what is behind this?”, “how would you describe it?”, “what happens on the other side”?, and also “what do you base this on?” The interviewer also asked to what extent the interviewee experienced this chatbot conversation as communication, and what this way of communicating tells them about the company. This last question helped us in answering RQ3 which focuses on company perceptions. After these questions, the interviewee was invited to use a second chatbot that we assigned to them, and subsequently they answered a shortened version of topic two for this second conversation.

The third and last topic tapped into the final evaluations regarding the implementation of chatbot conversations, and thus related to RQ3. The interviewer reminded interviewees of the expectations that they mentioned in the first topic, and asked about the extent in which the chatbot conversations during the interview fulfilled these expectations. Relatedly, the interviewer probed into how the conversation ideally should have been. The interviewer also asked about future uses of customer service chatbots: would they use them again, and if yes, what would they use them for, and in which types of moments. We also asked for what types of companies, and for what types of questions they think chatbots are appropriate.

Selection of Chatbots. We selected nine customer service chatbots that together showed variation on two dimensions: first, chatbots for profit as well as non-profit organizations were selected, and second, chatbots with humanlike characteristics versus chatbots with more robotlike characteristics were selected. We ensured that we had two humanlike chatbots for profit organizations, two humanlike chatbots for non-profit organizations, etc. All chatbots were available on the companies’ websites. For each chatbot, we prepared a scenario with a question that the chatbot was able to answer, and one that the chatbot was not able to answer. Prior to each interview, we assigned two chatbots to the interviewee based on their customer characteristics to ensure that they would use a chatbot of a company that they would contact in real life.

3.2 Sample

Research agency Ruigrok Netpanel coordinated the selection of the interviewees, with the help of ISO-certified agencies specialized in respondent recruitment. The interviewees received a monetary compensation for their participation, in line with the normal procedures of Ruigrok Netpanel. Interviewees were not allowed to have participated in qualitative research in the six months before the interview. A selection criterion was that the interviewee should have experience with contacting companies.

We ensured that the sample of interviewees (N=24) was varied in terms of gender (male N=12; female N=12), age (18-25 years N=5; 26-35 years N=15; 36-45 years

N=5; 46-65 years N=5; 65-78 years N=4), educational level (low N=8; middle N=8; high N=8) and household composition.

3.3 Data Analysis

First, the two authors working for the research agency wrote a research report (in Dutch, available upon request) in a way that is typically employed by research agencies. They summarized the main findings for each of the three research questions, and illustrated these with numerous quotes. Subsequently, the first author analyzed the interviews, taking the research report as a starting point. All interview transcripts were uploaded in the computer program Atlas.ti. In Atlas.ti, she conducted open coding, a procedure commonly used as the first step in the Grounded Theory Approach [4]. She read the interviews closely, line-by-line, and -for each research question- added codes. After going through the interviews, she wrote the current result section. The contents of the result section are in line with the research report, but the section does include some additional insights, such as the paragraph on the effort that customers need to put into communication with a customer service chatbot (RQ2).

4 Results

Based on the interviews, this section describes the communication journey of customers who wish to contact companies, and the evaluation of chatbot communication within this journey.

4.1 Customers' Prior Expectations of Communication with a Company

RQ1 asked about customers' expectations and moods prior to starting a conversation with a customer service chatbot. The starting points for interviewees' communication journeys with companies were -obviously- their questions or complaints. The interviewees expressed clear expectations of companies. Of utmost importance is getting their question answered or complaint addressed. This needs to be done fast (directly, within a few hours, for some questions a few days could be fine). They want to be taken seriously, and they want to be helped in a friendly manner. Sometimes it is needed that the customer is addressed with his/her name, that there is continuity over the course of several contacts, and in case of problems empathy is required. Interviewees expressed that sometimes they come into such conversations in a bad mood, even angry, for instance because they need an arrangement because they can not pay a bill; products or internet/phone connections do not work, or a package did not get delivered.

So what are the modes of communication through which these expectations can be fulfilled? And -importantly- are chatbots among these? Interviewees indicated that they first search for information online, and in case additional contact is necessary they rely on calling, e-mail or chat. In terms of preferred modes of communication, we can distinguish "phone callers" and "typers". "Phone callers" are interviewees who emphasized human contact, hearing someone's voice, being able to ask follow-up questions

and talking back and forth to come to an answer. Talking to a person makes it more direct and honest, and one immediately gets an answer. One can also express frustrations. The disadvantages of calling are the waiting times, that call center agents sometimes work with sale scripts, and that it is not available 24/7. “Typers” were the interviewees who preferred e-mailing or live chat (i.e., chatting with a human being). They indicated that they sometimes feel less comfortable with talking on the phone, and that they feel the need for having the information in writing for later referral. An additional advantage is multitasking: being able to send a mail or chat message while being at work. It is relevant to note that also being in contact with someone through live chat can be experienced as personal and having enough human touch. As interviewee 16 (female, 36 years, management assistant) said: “as long as it is a human being”.

Within this interview context of openly discussing how one prefers to contact companies, nine interviewees mentioned chatbots (and more mentioned them later in response to the chatbots in the interviews). Reasons for using chatbots were: finding information (comparing the function of chatbots to Google), being able to contact the company outside of office hours, and getting through to a live chat.

Several interviewees spontaneously described problems with chatbots, expressing mild frustration to clear anger. They expressed a lack of confidence in chatbots’ abilities. Chatbots do not understand their questions and keep asking “do I understand you correctly?”. Interviewee 16 described coming across chatbots when the live chat is closed in the evening: “and then you also have chatbots, but there you really do not get anywhere, so then I rather send an e-mail. Yeah, [with chatbots] you type a question, but they cannot give you a clear answer because it is such a robot of course”.

Based on their negative expectations with chatbot interactions in real-life, some interviewees already started the chat during the interview with “oh no, not a chatbot!”. For instance interviewee 7 (female, 29 years, cook) had already indicated her frustration with previous chatbots. Starting the chat, she exclaimed “this is clearly such a robot, I don’t like it”. Several interviewees expressed that normally they would avoid using chatbots, and rather wait “till Monday” to have a phone call or send an e-mail.

4.2 Customers’ Experiences during Chatbot Conversations

RQ2 asked about the most important characteristics of customers’ experiences with chatbot conversations, and asked specifically about their perceptions of the humanization of these chatbots and their language style. During the interviews, interviewees used two customer service chatbots (without us saying that these were chatbots). In the user experiences during the interviews, four elements stood out.

Prerequisite: Receiving Adequate Help. As expected, the main prerequisite for a positive evaluation of the conversations was that the interviewee felt that he/she had received adequate help. To exemplify such positive experience:

He [the chatbot] immediately gave a lot of information. As I just mentioned, that is what I really like. That I do not have to probe further. And even though it is digital, it is very friendly. It is more than enough information for me, so it really helped me. [...] I would give it a 10. Nothing is missing. (Interviewee 3: male, 19 years, student)

This interviewee said that this chat met his expectations, and that he prefers a robot “because you get the answers and you get them fast”. Even though this interviewee judges this chat as “perfect”, he does not see chatbot communication as sufficient enough to function as a stand-alone mode of communication. According to him, communication with a human is still required in case of problems, for instance when a package has not been delivered.

What interviewees particularly appreciated was that some chatbot conversations helped to filter information or that the chatbot provided links where the interviewee could find the information. As an interviewee positively noted: “Nice! Saves some searching”. Interviewee 16 said: “It was fine, clear, right information”. As mentioned, she had had negative experiences with chatbots before and had indicated that she would not use them. This chatbot conversation went fine, but did not erase her negative expectations “ok, this was fine, but I don’t know what would have happened if the question would have been more complicated”. In other words, despite of this positive experience, she was still expecting negative experiences to come.

In quite some cases, the chatbot did not meet the requirement of answering the question satisfactorily. The interviews clearly showed us how this leads to quite some frustration. Interviewee 1 (male, 64 years, human resources) almost started screaming:

I want to stop; I don’t get an answer. This was my first question, then I get this [shows answer]. Then I pose the second question, and I get: ‘thanks for your response, with your input I can improve myself’. I want to scream: ‘no, I was looking for something else!’ The biggest problem here is that I do not get an answer to my question, so I want to stop this [...] In my perception this just does not work: how precise can I be?

Effort. The interviews clearly revealed that it takes quite some effort for customers to use chatbots. This is partially related to the just described situation of not getting an answer, but there are also other ways in which customers are putting in effort when using chatbots.

Starting the chatbot conversation, some interviewees were wondering and estimating whether this chatbot would actually be able to answer their question. Some of them thought that for simple questions it may work, but not for more complicated questions. After considering this, they are making an effort to formulate the question in such way that they think the system will be able to understand it. Some say: “you need to use the right key words”. A complication here is that chatbots currently available do not all operate the same way: some indeed work best with key words as input, whereas others instruct the user to type full sentences like in human-human conversations.

It was quite common that, after typing the question, the chatbot did not come back with a relevant answer, and the interviewee needed to reformulate the question, sometimes several times. As the above-mentioned angry interviewee 1 expressed: “I already tried to ask the same question twice; now I would have to formulate it a third time; I am not going to do that”. In case there is no alternative for the company (i.e., the customers needs this specific company) and/or there is no other means of contacting the company (e.g. late at night), one feels convicted to keep trying to get an answer even when one does not feel like it. When in the end the result is still unsatisfactory, some

interviewees sighed: “see, this is what I mean, this all takes time, and now I would have to find another way of finding the information”. These efforts sometimes trigger prior negative experiences with chatbots. For instance interviewee 2 (male, 30 years, train driver), who did not mention chatbots during the first topic, said:

See, and then I am sitting at home and it annoys me. Then I think ‘guys!’ Often it does not work, or it is too slow, or indeed -that happened to me too- then it says ‘I do not understand it, can you formulate it differently’. Yeah, and that takes time, and you just don’t feel like that.

Simulation of Human Touch. An essential feature of chatbot communication is that it can appear humanlike due to anthropomorphic cues and its conversational nature. At the end of their interviews, two interviewees still thought they had been exchanging messages with human beings. Although interviewee 9 (female, 78, retired) still thought this, she did start doubting it during the course of the interview: “I HOPE it is a person!”. She was not completely out of touch with technological possibilities, but expressed concerns about her digital literacy related to her age: “I do have Instagram, to stay in touch with the children. But I don’t know whether young people understand how difficult it is for us [older people]”.

All other interviewees (sooner or later) were aware that they were using a chatbot, although they used many different names for it (a chatbot, chatbox, robot, computer, system, digital assistant, virtual assistant, algorithms, etc.). As interviewee 1 said: “I communicated with a computer, not with a lady Nina”. Also for these interviewees there was confusion sometimes, especially in the beginning of the conversations, and especially about whether it was a live chat (i.e. chatting with a human being) or a chatbot. Interviewees concluded that it was a chatbot predominantly with the help of the following cues: the responses came very fast, the responses were too similar, and their question was not understood. As one said: “You hope that it is a real agent, but the answers show that it is a robot”. Interviewee 4 (female, 40 years, operations manager) also expressed disappointment:

When you see the name ‘Billy’, you think it is a person. Also for me, although I do think that it is a virtual assistant. The first association is that it is a real human, and then the disappointment is bigger when you do not get anywhere [with getting the question answered]

Interviewees typically saw the conversation as a automated chat with a *simulated* human touch to it. The word “simulated” is important here. They expressed that with dialogical cues (“Hello, how can I help you today?”), a name, an icon or a picture of a human face, they make it seem that you talk to a real customer service agent. As interviewee 3 explained about the icon:

That is just a picture to make it seem as humanlike as possible [...] The same applies to the name probably; they want to make it look human. Giving it a name, an icon, in order for you to feel like you are communicating with a human being.

Opinions about this type of simulation varied. Some said they liked it and thought it was nice: “yeah, it is a bit more cozy”. Interviewee 8 (male, 46 years, male nurse) said:

Often you have this with such chat, that they have given it a name. I like that more. Then it looks more like ‘oh, you are not talking to a virtual assistant, you are talking to something with a name. Then it’s more personal.

Others said they do not care about this simulation. As interviewee 2 said: “Names I do not care about that much, but I do want to know: do I speak to a robot or... I do not even know what it is called.. an automated thing”.

And yet others were outrightly annoyed by the simulation and saw it as counterproductive. These interviewees saw it as a tricks and as a sign that the company was not taking them seriously. They became even more negative when their query was not answered. Interviewee 7 said about chatbot Iris (with name and cartoonlike icon): “Hahaha, that is Iris, robot Iris! They do this to make it look more personal, to give the impression that you are chatting with someone. I do not care that much; I don’t really look at it”. But later she turned much more negative: “Well, Iris, with the friendly smiling icon: probably they did this to make it look more personal. But they really made a mistake here. Because it is a very distant robot. And inefficient too”.

Language Style. One of the essential aspects of chatbot development is conversational design. Interviewees typically responded to our question about the chatbot’s language style by saying something about whether the style was formal versus informal. More specifically, they often mentioned the use of “u” versus “je”, which are the two nouns in Dutch to say “you” in a formal versus more informal way. Related to the communication expectations mentioned above, they also judged whether the conversation was friendly enough. In most cases, interviewees were fine with the style. As interviewee 3 noted:

For example the questions ‘how can I help you’, ‘let’s see what we have got’: just very clear, very friendly. Not too blunt, answers are nicely elaborate. Also important that in the end he asks ‘is there something else I can do for you?’, those kinds of things. Then you see that it comes across as very friendly.

However, some interviewees expressed their discontent with informal language. Interviewee 5 (female, 64 years, management assistant) said: “the use of ‘jij’ everywhere, I think it is too informal, but there is nothing we can do about it, we just have to accept it”. Sometimes such opinions are related to a specific company, for instance “the Douane [i.e., customs] should say ‘u’” (interviewee 6, female, 52 years, front office employee).

4.3 Overall Evaluations of Customer Service Chatbots and Company Perceptions

RQ3 tapped into customers’ final conclusions about under which conditions customer service chatbots can be implemented, and into the consequences of chatbot communication for customers’ company perceptions. On the positive side, interviewees concluded that chatbots have the potential to be helpful for certain purposes. Chatbots were considered to be good at helping with searching information. Some interviewees applauded the speed, and the 24/7 availability. Interviewee 2 said: “at least it points you in the right direction; now I do not have to search on the website. I fill out a word –

‘inboedelverzekering’ [‘contents insurance’]- I get some information, I can click on it, so that was fast”. This can be seen as a type of service, although they also say it is not always sufficient. Overall, interviewees were pretty clear that chatbots can be helpful for general and simple questions: it is perceived as a type of frequently asked questions, but presented in another way. If working well, it could also serve as a type of filtering, before being transferred to human agent. Some expressed understanding that the trend towards the use of chatbots and related technologies is inevitable, but that it needs to be improved.

However, interviewees were unified in that chatbots can not function as a stand-alone mode of communication, and that it can *not* replace human communication. As interviewee 5 said: “it will never meet the expectations, because it is not a human being, it will remain standard texts”. Chatbots can not answer all questions. Moreover, a human being can discuss with customers, can probe, and such conversation is more personal. Especially for personal and/or complex questions interviewees deem chatbots not useful because empathy is lacking. For instance interviewee 6 said that when she asks for money back, she wants to be comforted by a human being. Relatedly, interviewees mentioned that for certain companies and organizations, such as the police, hospitals and health care providers, chatbots are not suitable because these companies deal with personal situations. As interviewee 3 explained about contacting a doctor: “such issues most likely influence your body or your personal situation, so then I would prefer to have personal contact with a doctor versus via the chat”.

Company Perceptions. So what do these experiences tell the interviewees about the companies? On the positive end, some interviewees saw the companies as innovative. Some also saw it as a service for customers. However, interviewees typically thought that the implementation of chatbots was done more for the company itself than for the customers. As interviewee 8 said:

It looks they wanted to do it as cheap as possible. They did not pay attention to making it a bit nicer. Either they did not make an effort to find out how that could be done, or they did not come up with the right ideas to make it nicer.

A recurring thought in the interviews was that the primary reason for companies to use chatbots is to save costs. Their reasoning is that chatbots replace employees and that this saves money. As interviewee 1 expressed: “This is cost cutting, I know that, because that is also why I lost my last job”.

Some interviewees were also surprised about how bad the technology still is, especially when one had higher expectations of the level of innovation of a particular company. For instance, interviewee 7 had high expectations of the chatbot of a particular company, and was negatively surprised that this chatbot was not able to answer her question: “with the [name of another company], I kind of have an old-fashioned feeling, but with [name of company] you think: it is a very large company, everything will function well there”.

The effort that it takes customers to use the chatbot, and the frustration, can contribute to negative statements about a company. The above-mentioned angry interviewee 1: “what do you want to achieve with this mister [name company]? Maybe in their eyes it is customer friendly: ‘we help our customers’. Oh really? Come and use it yourself”.

5 Discussion

In the following sections, we will highlight the main findings for each of the three stages in the communication journey that customers have with companies. For each stage we will connect the findings to previous and future research, and outline the practical learnings. After this, we will describe some limitations of the current study and present additional suggestions for future research.

5.1 Customers' Prior Expectations of Communication with a Company

The current interview study showcases the expectations that customers have, and the moods that customers are in, when starting conversations with customer service chatbots. This is important because previous research taught us that the state that customers are in when they begin a chatbot conversation plays a role in how they experience the chat [9].

Our interviews reveal that customers' main priority is -obviously- to receive help with their customer query: their questions need to be answered; information needs to be provided, and their complaints need to be addressed. This has to be done fast, in a friendly manner. Some customers come into the conversation frustrated or angry. This may not only be because of the nature of their question or complaint, but -importantly- also because of previous interactions with chatbots (of other companies) that did not go well. For future experimental research, this implies that angry or frustrated participants need to be included in the sample, or that participants need to be brought in such state prior to the conversation.

In terms of practical recommendations, these findings imply that it is of utmost importance that developers take into account that there may be quite some resistance on the side of customers to overcome. Companies need to use the information they have about the states of their customers. Automated sentiment analysis will grow more sophisticated in the future and may make it possible to detect the mood that a specific customer is in when starting the chat. However, it needs to be kept in mind that interviewees expressed that in case of anger, frustration or fear (e.g. when not being able to pay the bills), empathy from fellow human beings is called for.

5.2 Customers' Experiences during Chatbot Conversations

In the four aspects of user experiences that we found, we see a similar hierarchy as described by Følstad and Skjuve [6]: for customers, receiving adequate help is much more important than the humanization of the chatbot or the language style. Our finding that users have to put in quite some effort seems to deviate slightly from findings from previous qualitative work [6]: our interviewees seem to express more frustrations and anger about chatbots' functioning. This disparity is most likely due to the fact that our sample included interviewees who were rather unwilling to use a chatbot, whereas the previous study [6] "only" included interviewees who actually had chosen to use a chatbot. Moreover, our finding is in line with other types of research that displayed quite

some issues and errors in the current generation of customer service chatbots [7, 11, 12].

With the first practical recommendation we echo previous work: the top priority should be to help customers in a sufficient and efficient way. Currently it still occurs too often that a customer does not find what he/she is looking for. The experience should be as seamless as possible to avoid (further) fuelling of frustrations. Placing a chatbot online early in its development may be useful for data collection and scaling the chatbot. However, looking at it through an individual customer's eyes, one needs to realize that this scaling approach can lead to quite some effort and frustration on the individual level.

Another issue is the simulation of human touch. Some users may indeed find it entertaining, or on an unconscious level it may trigger positive responses. However, it also needs to be realized that it can be problematic, both from a business perspective, as well as from an ethical perspective. The current interviews showed that humanization may also backfire: users can see it as an unwanted trick or as deception. Hadi's research implied that this is specifically the case when customers are angry when coming into the chat [9], and the current interview study illustrated that customers are indeed angry and frustrated at times. From an ethical perspective, it is problematic that some users trust the information because they incorrectly think they are communicating with a human being. This means that transparency is called for. It seems to become the social norm that a chatbot should be identified in the introduction section as a chatbot, or virtual agent, but the interviews illustrate that quite some users overlook such disclosure.

5.3 Customers' Overall Evaluations and Company Perceptions

In line with previous qualitative work [6], we found that some customers have specific ideas about what chatbots currently can and can not do. However, they still end up disappointed with the results. Thus, it is important that developers have a clear view of what the purpose of a specific chatbot is, and this should be communicated as clearly as possible to the user. Providing users with more guidance regarding the queries a chatbot can help with may be useful.

In terms of company perceptions, the interviews showed that the implementation of chatbots is currently quite a risky adventure for companies. Although some interviewees related the chatbot to the innovative character of the organization, it was common that they saw it as a result of cost-cutting. Also, interviewees typically thought the implementation of chatbots was done more for the company itself than for the customers. In terms of an organization's image, it is relevant to consider whether these are the associations that one would want to trigger.

5.4 Limitations and Future Research

The focal point of the current analysis was the communication journey that customers have with a company, and a chatbot's place in it. This means that there are specific elements in the user experiences that call for more in-depth analyses; in particular the

responses to the anthropomorphic cues require a more detailed look. The link between certain chatbot perceptions (e.g., seeing it as a search engine versus expecting answers to more complex questions) and ways of communicating with the chatbot (e.g. using only key words, or typing natural sentences) should also be explored.

By definition, after a qualitative study, quantitative follow-up studies are required. Surveys and experiments are needed to further investigate the relations between mood states prior to the chatbot conversation, and subsequent user experiences. Also the role of humanization within the overall user experience needs to be disentangled more fully. This work should also delve into the relations between demographics (such as age, gender, and educational level) and chatbot perceptions [1].

For all research into user experiences of chatbot communication goes that the findings are bound to a particular point in time, due to continuous improvements and AI developments. However, we do think that some of the experiences outlined in the current study will remain important for research and practice for quite some time.

Acknowledgements

This study was partially funded by *Logeion*, the Dutch association for communication professionals.

References

1. Araujo, T., ter Hoeven, C., van Zoonen, W.: Automated 1-2-1 communication. In: SWOCC, vol. 77. Stichting Wetenschappelijk Onderzoek Commerciële Communicatie (SWOCC), Amsterdam (2019)
2. Araujo, T.: Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Comput. Hum. Behav.* **85**, 183-189 (2018). <https://doi.org/10.1016/j.chb.2018.03.051>
3. Brandtzaeg, P.B., Følstad, A.: Why people use chatbots. In: Kompatsiaris, I., et al. (eds.) INSCI 2017. LNCS, vol. 10673, pp. 377-392. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70284-1_30
4. Charmaz, K.: *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage, London (2006)
5. Følstad, A., Nordheim, C.B., Bjørkli, C.A.: What makes users trust a chatbot for customer service? An exploratory interview study. *International Conference on Internet Science*, pp. 194-208. Springer (2018). https://doi.org/10.1007/978-3-030-01437-7_16
6. Følstad, A., Skjuve, M.: Chatbots for customer service: User experience and motivation. *Proceedings of the 1st International Conference on Conversational User Interfaces (CUI 2019)*, pp. 1-9. ACM, New York (2019). <https://doi.org/10.1145/3342775.3342784>
7. Gnewuch, U., Morana, S., Maedche, A.: Towards designing cooperative and social conversational agents for customer service. *Proceedings of the 38th International Conference on Information Systems (ICIS) (2017)*
8. Go, E., Sundar, S.S.: Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Comput. Hum. Behav.* **97**, 304-316 (2019). <https://doi.org/10.1016/j.chb.2019.01.020>
9. Hadi, R.: When humanizing customer service chatbots might backfire. *NIM Marketing Intelligence Review* **11**(2), 30-35 (2019)

10. Ischen, C., Araujo, T., Voorveld, H., van Noort, G., Smit, E.: Privacy concerns in chatbot interactions. In: Følstad, A., et al. (eds.) Chatbot Research and Design: Third International Workshop, CONVERSATIONS 2019. Lecture Notes in Computer Science; Vol. 11970, pp. 34-48. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-39540-7_3
11. Rese, A., Ganster, L., Baier, D.: Chatbots in retailers' customer communication: How to measure their acceptance? *Journal of Retailing and Consumer Services* **56**, 102176 (2020). <https://doi.org/10.1016/j.jretconser.2020.102176>
12. Sheehan, B., Jin, H.S., Gottlieb, U.: Customer service chatbots: Anthropomorphism and adoption. *Journal of Business Research*, **115**, 14-24 (2020). <https://doi.org/10.1016/j.jbusres.2020.04.030>
13. van der Goot, M.J., Pilgrim, T. : Exploring age differences in motivations for and acceptance of chatbot communication in a customer service context. In: Følstad, A., et al. (eds.). Chatbot Research and Design: Third International Workshop, CONVERSATIONS 2019. Lecture Notes in Computer Science; Vol. 11970, pp. 173-186. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-39540-7_12