The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users

Conference Paper · July 2023

DOI: 10.1145/3571884.3597144

3 authors:

Marita Bjaaland Skjuve
SINTEF
25 PUBLICATIONS **701** CITATIONS
SEE PROFILE

Asbjørn Følstad
SINTEF
133 PUBLICATIONS **5,369** CITATIONS
SEE PROFILE

Petter Bae Brandtzaeg
University of Oslo
102 PUBLICATIONS **5,417** CITATIONS
SEE PROFILE

The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users

Marita Skjuve
SINTEF Digital
marita.skjuve@sintef.no

Asbjørn Følstad
SINTEF Digital
Asbjorn.Folstad@sintef.no

Petter Bae Brandtzaeg
SINTEF Digital, University of Oslo, Department of Media and Communication
p.b.brandtzag@media.uio.no

ABSTRACT

The launch of ChatGPT has attracted significant attention and showcased the potentially game-changing capabilities of conversational AI. These capabilities, and lack of user research, highlight the need to investigate how users experience interactions with conversational AIs like ChatGPT. Therefore, we conducted a questionnaire study with ChatGPT users (N=194), inquiring about their good and poor experiences with ChatGPT. The user reports were analyzed by a thematic analysis and systematized through a pragmatic-hedonic framework. Our results demonstrate how user experience is influenced by pragmatic attributes such as ChatGPT providing useful and detailed information and easing work- or school-related tasks. Additionally, user experience is impacted by hedonic attributes, such as entertainment and creative interactions, and interactions leaving the user impressed or surprised. Our study underscores that user experience concerning conversational AI like ChatGPT is assessed by useful and productive interactions even in early phase of uptake, suggesting the importance of pragmatic attributes.

CCS CONCEPTS

• **Human-centered computing**; • **Empirical studies in HCI**;

KEYWORDS

Conversational AI, ChatGPT, User experience, Pragmatic-hedonic framework

ACM Reference Format:
Marita Skjuve, Asbjørn Følstad, and Petter Bae Brandtzaeg. 2023. The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users. In *ACM conference on Conversational User Interfaces (CUI '23), July 19–21, 2023, Eindhoven, Netherlands.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3571884.3597144

1 INTRODUCTION

Open AI launched ChatGPT on November 30, 2022. Immediately following its debut, ChatGPT was characterized as a potential game-changer [1] with record-breaking growth in user base. As of January 2023, the service had 13 million daily users [2], with numbers rapidly increasing.

CUI '23, July 19–21, 2023, Eindhoven, Netherlands

ACM ISBN 979-8-4007-0014-9/23/07.
https://doi.org/10.1145/3571884.3597144

ChatGPT is a conversational user interface to the large language model (LLM) Generative Pre-trained Transformer 3 (GPT-3). However, the underlying language model will keep changing; it is now at GPT-4 for paid version. ChatGPT is pre-trained on a large dataset and fine-tuned through supervised and reinforcement learning. ChatGPT can carry out various tasks, including question answering, text generation and summarization, text improvement, and code production [3, 4]. ChatGPT differs from previous conversational user interfaces, as it is able to generate relevant and specific responses for a broad range of topics and domains, due to its basis in a state-of-the-art large LLM. In the following we use the term use conversational AI to distinguish conversational user interfaces based on LLMs – of which ChatGPT is currently the primary example – from previous types of conversational user interfaces.

Conversational AI is predicted to become the next generation of digital services for automated assistance or support [5]. This next generation of conversational AI, with a broad range of capabilities and sophisticated natural language capacity, will likely substantially impact how we interact with technology [6].

The predicted impact of conversational AI, such as ChatGPT, makes it important to gather initial knowledge of how users experience this technology. With ChatGPT, users can accomplish a wide range of tasks using a single user interface, suggestive of the disruptive potential of conversational AI on the way we acquire, process, and generate knowledge. An important concern is that the rapid and ongoing advancements in conversational AI can be attributed to a massive technological push. This may be particularly true with the launch of ChatGPT and the AI race between Google, Microsoft, and other tech companies. So far, little systematic knowledge, beyond news media stories, is available on how ChatGPT users experience. It is essential to gather such knowledge at an early uptake stage - not only to guide the development of conversational AI but to enable future studies to investigate how the user experience of conversational AI evolves. Thus, we asked the following research question: What characterizes the user experience of interactions with ChatGPT?

In response to this question, we conducted a qualitative exploration involving ChatGPT users in a questionnaire study. Here, we asked for examples of users' good and poor experiences. These user examples were examined using thematic analysis and organized based on the pragmatic-hedonic theoretical framework.

Our findings provide valuable insight into key pragmatic and hedonic attributes influencing user experience of ChatGPT and, by extension, conversational AI. These results will hopefully enable the industry and research community to better understand, prepare, and design for the uptake of conversational AI.

2 BACKGROUND

2.1 User Experience – Theoretical Basis

User experience is a prioritized research area for conversational user interfaces [7]. This has resulted in a broad range of studies on how users perceive such interfaces and why they engage with them. User experience is defined as a "user's perceptions and responses that result from the use and/or anticipated use of a system, product or service" [8]. A range of theoretical models [9] and measurement instruments [10] have been proposed to analyze and investigate different dimensions and attributes of user experience [11, 12]. Previous research has found the pragmatic-hedonic framework to be a helpful theoretical perspective for analyzing the diverse range of user experiences associated with conversational user interfaces [13]. Pragmatic qualities refer to user experience attributes reflecting usefulness and usability, while hedonic qualities refer to the system's capacity to appeal to a user's desire or sense of pleasure (e.g., entertaining, engaging, interesting) through stimulating, evocative experiences in line with the user's identity [14].

A questionnaire study exploring user experience in the context of chatbots [13] found the pragmatic-hedonic framework to cover user experience attributes concerning assistance, information, entertainment, and inspiration. In addition, user experience attributes were observed beyond the pragmatic-hedonic framework, such as the social and humanlike characteristics of the interaction.

We have in this study chosen to apply the pragmatic-hedonic framework to study ChatGPT, since it has been successfully applied in studying user experience of chatbots [13]. Furthermore, the distinction between pragmatic and hedonic dimensions of user experience is found in most models of user experience [43], suggesting the general applicability of this framework.

2.2 User Experience – Research on Conversational User Interfaces

Conversational AI, understood as general-purpose conversational user interfaces to LLMs, distinguishes themselves from traditional conversational user interfaces in that they are driven by LLMs which enable flexible and comprehensive interactions across a wide range of uses cases. As such, Conversational AI represents a novel paradigm and may produce new user experiences.

However, previous research on user experience of conversational user interfaces is likely to have substantial relevance also for conversational AI. This research has focused on two broad application areas [15]: task-oriented solutions, such as those found in customer service and assistants, and socially oriented solutions, such as companion chatbots and solutions for well-being and mental health.

For task-oriented conversational user interfaces, pragmatic quality attributes such as efficiency and help-giving strongly determine user experience [16]. Task-oriented conversational user interfaces typically appeal to goal-oriented users with concrete needs [17]. For instance, Mogaji, Balakrishnan, Nwoba and Nguyen [18] reported on an interview study with conversational user interfaces in the banking domain. User experience was influenced by access and efficiency of the provided support. Likewise, Oesterreich, Anton, Schuir, Brehm and Teuteberg [17] accentuated the need for agents to avoid information overload and provide conversational repair in case of misinterpretation or erroneous information in their design principles for voice agents in customer service. Failing to provide sufficient pragmatic quality in task-oriented solutions may be detrimental for user experience [19]. van der Goot, Hafkamp and Dankfort [20] found negative implications of limited language understanding and knowledge resources in customer service chatbots. Likewise, Luger and Sellen [21], in an interview study with voice agent users, found them to be frustrated when agents could not perform requested tasks and showed limitations in natural language understanding.

For socially oriented conversational user interfaces, a different set of user experience dimensions and attributes is important [22]. For instance, Brandtzaeg, Skjuve, Dysthe and Følstad [23], in a study of youth's experiences with the therapeutical chatbot Woebot, found that participants enjoyed the perceived psychological benefits from their interaction, such as enhanced self-reflection and social support. Emotional and social dimensions of user experience have been identified as important also for companion chatbots [22, 24]. For example, whether the chatbot is perceived as empathic [25, 26] and adhering to social rules, such as turn-taking and reciprocation [27].

Lastly, factors outside the pragmatic-hedonic framework may impact the user experience. In particular, trust strongly impact user experience for task-oriented [28] and socially oriented [29] conversational user interfaces. Furthermore, the human likeness of conversational user interfaces has been identified as potentially important for user experience for task-oriented [30] as well as socially oriented [24] solutions.

The above research may serve as a benchmark for exploring user experience for new conversational AI, such as ChatGPT. However, given the substantial variation in what constitutes relevant user experience dimensions for different types of conversational user interfaces, explorations are needed to identify how the user experience of conversational AI may be characterized.

3 METHOD

3.1 Study Design

To capture the breadth of user experience with ChatGPT, we applied a questionnaire-based approach. We chose this over other qualitative methods such as user interviews, observation, or think-aloud techniques as our goal was to engage with a broad range of users to explore variations and convergences in use. In the questionnaire, we asked the participants to elaborate on good and poor experiences with ChatGPT (open-ended). We also asked about age, gender, education level, and country of residence.

Table 1 summarizes the questionnaire items. The results from Q1-Q4 and Q7-Q8 are presented in this paper. The participant responses to Q5–Q6 concerned user motivations rather than experiences. These are therefore not considered in this paper but will be reported on in a separate publication on motivation to use conversational AI.

3.2 Recruitment and Data Collection

We targeted ChatGPT users who were fluent in English. To access the participants, we used Prolific, which is an "established platform for online subject recruitment which explicitly caters to researchers"

Table 1: Overview of questionnaire items

Questionnaire items	
Q1	In your own words, what is ChatGPT?
Q2	When did you first use ChatGPT?
Q3	How often do you use ChatGPT?
Q4	In which contexts do you make use of ChatGPT?
Q5	Why do you use ChatGPT?
Q6	What is your main reason for using ChatGPT?
Q7	Tell about one really good experience you have had using ChatGPT
Q8	Tell about one really poor experience you have had with ChatGPT
Q9	How could ChatGPT be improved to serve you better?

Table 2: Example responses to questionnaire item Q7 – Q8

Response	Words
Concerning good experience (Q7): I once had a DIY problem. Basically I was unsure about the causes of a drop in pressure in my home's heating system. Rather than call a plumber, I asked chatGPT and it offered a surprisingly detailed and nuanced overview.	40
Concerning good experience (Q7): It was helpful in completing a email which needed to be written. This allowed me to send a more professional sounding email	22
Concerning poor experience (Q8): The tool confidently states things which are not true, which is worrying. It also doesn't cite sources, because it's not actually using sources, it's predicting words.	26
Concerning poor experience (Q8): In some cases, ChatGPT isn't able to give a response for several reasons so I can't get all my information from just ChatGPT alone.	24

[31]. Data collection was conducted January 23-25, 2023, and the participants received 1,5 GBP as an compensation upon finishing the questionnaire.

To check that participants did not mistake ChatGPT for something else (e.g., a customer service conversational user interface), we provided a clear description of ChatGPT and asked the participants to describe in their own words what ChatGPT was. We examined every response to identify participants who had mistaken ChatGPT for a different entity. We filtered out 85 respondents who had insufficient or no experience with ChatGPT. In addition, four participants were filtered out due to incomplete responses.

The final sample consisted of 194 participants, where 69 identified as female and 123 as male (two opted not to report their gender). The average age was 34 years (SD = 11, range 18 - 70). Most participants reported having higher education (163; 84%), while 16% (31) reported high school or secondary school as their highest level of education. The participants were located in the UK (163; 84%), the US (19; 10%), and Ireland (11; 6%). One did not disclose country of residence. Average number of words in answers to Q7 was 29 (range 2 – 88 words), while for Q8 the average number of words was 23 (range 1 – 73 words). Examples of responses are provided in Table 2.

The study was conducted following informed consent. The data collection was fully anonymous, and participants were informed they could withdraw anytime.

3.3 Data Analysis

We applied thematic analysis following Braun and Clark [32] to the open-ended responses. Here, we developed codes that reflected the theme(s) expressed in each response. A response could be associated with one or more codes. All codes were merged and combined to establish common themes across the dataset. Themes were then mapped to the dimensions of the pragmatic-hedonic framework. Themes that did not fit this framework were placed in a separate group named other attributes.

The first author conducted the initial coding of all responses. To ensure quality in the analysis, the first and second authors met and discussed the emerging themes and codes at regular intervals during the coding process. When the entire dataset was coded, the second author checked the analysis by going through all the codes. Disagreements and uncertainties were resolved through discussion.

4 RESULTS

4.1 Participant ChatGPT Usage

The participants were typically early users of ChatGPT. Most (102; 53%) reported using the service since November or December 2022. Another 40% (78) had started using ChatGPT in January 2023. Only a small number of participants (14; 7%) started using ChatGPT the week of the study.

The participants reported using ChatGPT fairly frequently. Over half reported using ChatGPT weekly or more (113; 58%). The rest used ChatGPT several times a month (69; 36%) or rarely (12; 6%).

Most participants estimated having used ChatGPT 20 times or less (150; 77%). The rest stated to have used it more than 20 times.

The participants typically reported using ChatGPT in their personal life (160; 82%). Some also reported using ChatGPT for work (74; 38%) or school-related purposes (32; 16%). Each participant could indicate multiple answer alternatives.

4.2 How do the Participants Understand ChatGPT?

As the first question in the questionnaire, we asked the participants to formulate in their own words what ChatGPT is. The main purpose of this question was to check the participant's understanding of ChatGPT. However, the thematic analysis of their reports provided interesting insights from the perspectives and sentiments reflected in the reports.

The participant reports suggested that ChatGPT was mainly understood from three perspectives: its (1) technical functionality, (2) uses and purposes, and (3) interaction capabilities. Typically, user reports reflected more than one of these perspectives (147; 76%), potentially indicating that ChatGPT may hold several interpretations for a user.

Perspective 1: Technical functionality (185; 95%). Nearly all indicated that they understood ChatGPT in part from a technical perspective, and described ChatGPT as an AI, chatbot, machine learning tool, natural language processing tool, or search engine. Some also described it as trained on large datasets or holding capabilities for learning.

> "ChatGPT is a chatbot launched with the help of an AI model to transfer learning with both supervised and reinforcement learning techniques." (ID196)

Perspective 2: Use and purpose (124; 64%). Most participants also described ChatGPT in terms of what it could do for them. ChatGPT was often described as answering questions, generating content, performing tasks, or participating in creative work. Some also mentioned that it could be used for research, education or to improve their understanding of topics.

> "ChatGPT is an advanced AI based chatbot that is used primarily for research and education." (ID149)

Perspective 3: Interaction capabilities (55; 28%). Some participants reported ChatGPT's interaction capabilities as a defining feature. These participants described ChatGPT as a tool that communicates in a conversational manner. Some also mentioned that it communicates like a human or mimics human interaction.

> "A chat bot, an openly accessible AI that was designed to provide human like-answers and humanlike interaction via a chat and learn and improve itself through those interactions." (ID72)

The sentiment of participant reports. A substantial part of the participant reports reflected a positive sentiment (50; 26%) by, for example, including descriptions noting that ChatGPT was able to answer any question or provide very intelligent answers.

> "ChatGPT is an AI chat robot which can answer questions, create poems etc based on whatever you ask it and will even write in Scots as well as English." (ID211)

Negative sentiment was only reflected in very few participant reports (2; 1%). One pointed out that ChatGPT was not performing as hoped, and, as demonstrated by the quote below, ChatGPT is inclined to lie.

> "An AI chat bot, which is pretty good at the moment, but has a tendency to lie through it's virtual teeth." (ID96)

4.3 What Characterizes Good User Experiences?

Most participants (152; 78%) reported good experiences with ChatGPT that occurred within a specific context - for example, in the context of daily life (23; 12%), knowledge work (28; 14%), software development (21; 11%), creative activities (30; 15%), education (9; 5%), or health (8; 4%). Some reported on specific interactions that did not occur in a specific context (41; 21%), such as getting information through ChatGPT. Four did not have a good experience.

Our analysis revealed three pragmatic attributes and three hedonic attributes positively affecting user experience. We also identified two attributes outside the pragmatic-hedonic framework: pleasant communication style and psychological impact (see Table 3). Note that each participant report could be coded as reflecting more than one user experience attribute.

Pragmatic attributes: Practical outcomes (77; 40%). Positive interactions were often reported to encompass interactions that resulted in a specific practical outcome that the participants valued. For some, this concerned efficiency, i.e., that ChatGPT was found to answer questions quickly or that output could enable faster task execution, as seen in the following quote:

> "I couldn't figure out how to write a formula in excel for work and I needed it done before the end of the day. I typed it into ChatGPT and it worked perfectly. This saved me so much time (ID65)"

For others, positive practical outcomes were linked to ChatGPT enhancing the quality of their work. Quality enhancements were often reported for written work, for example, when the participants had to write essays for school, draft a CV, or write an email. The latter is exemplified below.

> "I had to email my boss and struggled to explain what I wanted to say to him. Chat GPT helped me word it better." (ID90)

Some described good pragmatic outcomes for situations where ChatGPT enhanced their understanding of a topic. ChatGPT was reported to aid their understanding in various contexts, from solving crosswords to helping the participants conclude on a health problem.

> "A good Ie I had with ChatG"T is'when I told it the symptoms I was having. I realized that I have IBS from this, so it was pretty helpful." (ID47)

Finally, several participants experienced concrete problems or issues that ChatGPT had helped with or solved, as depicted in the following quote:

> "One project was giving me some real difficulty as I was not very familiar with the subject. I used ChatGPT to initially find some information but wound up using

Table 3: Identified attributes of good user experience

Attributes	Description	n; %
	PRAGMATIC ATTRIBUTES	
Practical outcome	ChatGPT enables efficiency and enhances quality of work.	77; 40%
Relevant and useful output	ChatGPT provides relevant and useful output.	33; 17%
Accurate or detailed output	ChatGPT offers output that is accurate, concrete, and detailed.	32; 16%
	HEDONIC ATTRIBUTES	
Surprising or impressive	ChatGPT is seen as exceeding expectations, impressive, or superior compared to existing solutions.	62; 32 %
Creative impact	ChatGPT supports creative activities such as story writing.	29; 15%
Entertaining interactions	Interactions with ChatGPT are perceived as entertaining.	21; 11%
	OTHER ATTRIBUTES	
Pleasant communication style	ChatGPT is seen as humanlike or intelligent, with simple language and good structure and flow.	33; 17 %
Psychological impact	ChatGPT lowers anxiety, reduces loneliness, or enables social recognition	6; 3%

all the information put forth by ChatGPT to complete my project." (ID144)

Pragmatic attributes: Relevant and useful output (33; 17%). Several participants reported positive experiences due to ChatGPTs' relevant or useful output. This experiential attribute was for various contexts, from factual knowledge to support to interactions about daily life. Some described the relevance and usefulness of ChatGPT in general, as seen below.

> "I have used it to query a topic. The answer was relevant. It assisted my understanding." (ID139)

Unlike the theme "practical outcome," where the participants stated that output from ChatGPT was used to solve a concrete problem, reports coded as "relevant and useful" do not indicate output being used – just that it was perceived as useful or relevant. A few also expressed that they appreciated interactions where ChatGPT offered output that they could utilize as a foundation and expand upon. As such, the beneficial output does not necessarily need to be useful in its initial form.

> "I asked chat gpt about an essay question and it provided a few hundred words that I could use as a starting point. This was very time saving." (ID13)

Pragmatic attributes; Accurate or detailed output (32; 16%). Some participants described the quality of the information provided by ChatGPT. They appreciated ChatGPT's ability to offer specific, concrete, and accurate information. Often, they noted that such accuracy was unexpected:

> "The coworker who Introduced me to It guided me through using it and we decided to see if we could get it to tell us how to create meth. It actually went into the chemistry and gave us detailed and seemingly accurate answers which surprised me." (ID200)

Others reported on the level of detail in ChatGPT's responses, with many positive experiences characterized by ChatGPT's ability to provide comprehensive descriptions of a phenomenon.

> "I used it to get information on the top 3 biggest tech releases at CES this year. It came back with a long and detailed set of information on the products." (ID11)

Hedonic attributes: Surprising or impressive (62; 32%). Several participants would report good experiences due to ChatGPT exceeding their expectations or leaving them impressed. Most did not include an explicit reason for this, but a few did. As exemplified in the quote below, the participant was pleasantly surprised by the level of detail in the answer from ChatGPT.

> "I used chatgpt to generate a movie title and synopsis based on my criteria and in the style of my favorite movies, and I was very impressed with the results. I was impressed that it was able to provide justification for my criteria and explain in what way it would be similar to the movies I listed." (ID78)

Some participants also noted that ChatGPT performed better than existing solutions, such as Google or other chatbots.

Hedonic attributes: Creative impact (29; 15%). Some participants reported positive experiences with creative implications. The participants described how the interactions with ChatGPT could facilitate inspiration, idea generation, or support creative activities.

> "I asked it to write a short horror story in the style of Stephen King. It was interesting to see the story be written in real time and, although it wasn't like a story written by a professional author, it was a fun read." (ID72)

Others described how they enjoyed ChatGPTs ability to provide detailed stories, help the participants see a creative problem from different angles, or help solve a concrete creative problem, as demonstrated below:

> "It gave me a very good Idea for a "Itua'Ion between two of my original characters that I had added to the canon. This helped my writer's block and creativity." (ID66)

Interestingly, one participant described it as a good experience that he would gain social recognition by posting ChatGPT's creative output on forums.

> "As previously referenced, I asked it to design a Hearthstone card. It came up with a pretty cool dragon

> creature, and I shared it on the subreddit, which got some upvotes and replies." (ID187)

Hedonic attributes: Entertaining interactions (21; 11%). Some participants reported good experiences that had an entertaining aspect. These participants would typically carry out creative activities or indicate exploring ChatGPT's capabilities by having it make up poems or stories which were seen as entertaining:

> "I made a story about a kid who pooped his pants but it smelled like Indy 500 burnouts. Here I am almost 40 years old, in the office with tears in my eyes, because I am laughing so hard at the stupidest story this AI made up." (ID174)

A few also described enjoyment from ChatGPT interactions due to unexpected output in response to their prompts:

> "Aside from using ChatGPT for work, I had a great time having it write short stories. My favorite was about a bear finding a place to poop in the woods. Oddly, the poop bit was not part of the prompt." (ID86)

Other attributes: Pleasant communication style (33; 17%). Some participants liked how ChatGPT used simple language that was easy to understand or that it had a good structure to the conversations - as the participant below explains:

> "I love how the answer is typed out while you read, so a long answer doesn't hit you or even put you off. I also like the structure and extensiveness of answers." (ID20)

A few explained how the output was very humanlike or mentioned liking how interactive ChatGPT is, which makes it possible to get more information by asking follow-up questions.

> "it helped me understand the difference between irony and satire [. . .] chatgpt responds to follow up questions based on what you already asked so this is what helped me to understand." (ID12)

Other attributes: Psychological impact (6; 3). A few participants described good experiences due to the psychological implications of their interactions with ChatGPT. One reported interaction with ChatGPT that enabled self-reflection. Others noted feeling less down or lonely following interactions with ChatGPT.

> "I talked to it when I was in a really depressed mood, feeling at my absolute worst, and was able to get a lot off my chest that I can't tell other people. It responded quite genuinely and helped me feel better. This made me realize AI chatting could have a purpose in day-to-day life [. . .]." (ID4)

4.4 What Characterizes Poor User Experiences?

Several participants recalled having poor user experiences with ChatGPT, though to a lesser extent than good user experiences. Specifically, 45; 23% stated no poor experiences with ChatGPT, while 26; 13% reported no poor experience that stood out but still commented on things they did not like.

The reported poor user experiences were mostly found to occur in the same types of contexts as the good experiences, including daily life (6; 3%), knowledge work (2; 1%), software development (11; 6%), creativity (8; 4%), and education (2; 1%). Some reported on specific interactions that did not occur in a specific context (40; 21%), such as attempting getting information through ChatGPT.

We identified six pragmatic and one hedonic attribute that impacted user experience in a negative way (Table 4). We also identified two attributes that did not seem to fit within the traditional pragmatic/hedonic framework: repetitive and problematic communication style.

Pragmatic attributes: Irrelevant and useless output (39; 20%). Negative user experiences were frequently linked to ChatGPT's failure to provide relevant or useful responses, resulting in output unrelated to their original queries.

> "I asked for detailed information of a famous figure and it returned gibberish." (ID110)

Some noted that ChatGPT suggested wrong solutions, for instance, when doing math, or solutions that were not functional, for instance, when proposing code – as seen below:

> "ChatGPT provides codes that is not valid. Main reason for invalidity is it provides some hypothetical functions that the package does not have." (ID80)

Lastly, some participants noted that the ChatGPT offered responses that were too general or vague and thus not useful:

> "Some of the answers are fairly vague and too generalized; when in reality the questions are much bigger and complicated and should be treated as such." (ID84)

Pragmatic attributes: Challenging to formulate requests (28; 14 %). Some participants reported interactions where ChatGPT struggled to understand. A few emphasized that ChatGPT failed at understanding even when they tried to instruct it to help. Challenges in formulating requests were typically seen concerning highly specific tasks, such as solving a calculus issue in a specific way:

> "A poor experience I had is when I asked ChatGBT to solve a pretty simple maths equation. It kept getting it wrong and couldn't get it right even when I told it how it's done." (ID47)

Some participants also described poor user experiences in cases where they needed to communicate in a distinct way or had to formulate several prompts to get a relevant output.

> "Sometimes you have to prompt it to give you the answer that you're looking for." (ID37)

Pragmatic attributes: Technical issues (25; 13%). Some participants reported negative experiences concerning technical issues. In most cases, this was about ChatGPT being unavailable due to high demands and server overload. Some also experienced error messages that interfered with their interactions.

> "There was a bad experience when I tried to use chatgpt however it said its at capacity and I only use it for emergencies so I was disappointed when it said that therefore I had to comeback later." (ID145)

Pragmatic attributes: Inability to answer (24; 12%). Another negative experience was users that described not receiving any reply from ChatGPT. In some cases, ChatGPT did not know how to answer or refused to answer - as depicted by the quote below:

Table 4: Identified attributes of poor user experience

Attributes	Description	n; %
	PRAGMATIC ATTRIBUTES	
Irrelevant and useless output	ChatGPT provides irrelevant, off-topic, or vague information.	39; 20%
Challenging to formulate requests	The user reports struggling with getting ChatGPT to give the desired output.	28; 14%
Technical issues	ChatGPT is unavailable due to server overload or other forms of technical issues.	25; 13%
Inability to answer	ChatGPT does not want to answer the request due to policies or because it only has limited information in its database.	24; 12%
Trust issues	ChatGPT presents misinformation or biased views.	23; 11%
	HEDONIC ATTRIBUTES	
Not supporting creative activity	ChatGPT fails at generating output that supports creative activities, such as being poor at story writing.	9; 5%
	OTHER ATTRIBUTES	
Repetitive	ChatGPT repeats the outputs and keeps presenting similar information.	15; 8%
Problematic communication style	ChatGPT communicates in a way that feels fake or not humanlike. ChatGPT can also be rude.	9; 5%

> "I asked it whether it thought AIs should govern humans. It offered a noncommital answer along the lines that it's not in a position to decide on such an issue." (ID113)

Other participants would highlight issues caused by ChatGPT having limited or outdated information in its database. Due to this, ChatGPT could not provide an answer to the participants' requests.

> "Whenever I've asked it about local based news or history. I can't remember specifically what ChatGPT responded but it was along the lines of not having the capabilities to formulate an answer." (ID158)

Pragmatic attributes: Trust issues (23; 11%). Several participants described negative experiences due to trust-related issues. Some explicitly pointed out ChatGPT's tendency to deliver inaccurate information or make up things (e.g., hallucinate) yet present it as correct or adequate. This is exemplified in the participant's quote below:

> "It frequently makes stuff up and can confidently spout BS. I asked it for the definition of terms that I made up and it provided a definition that seemed convincing but was not real. I think OpenAI is going to have a difficult time preventing ChatGPT from providing inaccurate information." (ID164)

Although this attribute somewhat overlaps with 'irrelevant and useless output' described above, we coded trust issues for instances where participants expressed it as problematic that output was presented as truthful information. Some participants explicitly said they did not trust ChatGPT and stressed the importance of double-checking the answers it provided. A few participants also experienced that ChatGPT contradicted itself or that it offered biased information:

> "I was asking about dating trends and some relevant advice which ended up contradictory to my values, however popular in certain circles." (ID92)

Hedonic attributes: Not supporting creative activity (9; 5%). A few participants reported negative interaction because ChatGPT could not support creative activities. This was often seen in relation to wanting ChatGPT to write a story or a poem in a specific style, as seen below.

> "all the iambic pentameter had more than 10 syllables, they should only have 10. When i asked it to check if it had 10 syllables it said sorry your right. But still couldnt give me one with 10 syllables in." (ID207)

Other: Repetitiveness (15; 8%). Some participants reported interactions where ChatGPT kept repeating itself, even when they tried to get it to generate other replies.

> "I had difficulty in getting good book recommendations from chatgpt because it kept giving the same few books for each question which made me think it's database was limited. It also gave me the wrong information for one of the books." (ID78)

Other: Problematic communication style (9; 5%). A few participants described that ChatGPT came off as fake or that they could sense that they were interacting with an artifact aimed to be humanlike. As indicated by the participants below, it did not necessarily cause a bad experience but was perhaps disappointing.

> "Nothing was really a poor experience with it. You can clearly tell you're talking with a robot though. It's no more 'life-like' that other chatbots." (ID140)

One participant reported an experience where ChatGPT presented disturbing information, while another found ChatGPT to be rude. As seen below, the participant found themselves "told off" because ChatGPT did not like their request:

> "Told me off for trying to make it write fanfiction (not even explicit stuff just fun!)." (ID154)

5 DISCUSSION

Our study shows the significance of pragmatic attributes for the user experience. Even during initial adoption, useful and productive outcomes and interactions are apricated. These results also demonstrate the importance of hedonic attributes where good user experiences arise when ChatGPT supports participants in creative activity or impresses them.

We will in the next section address how ChatGPT is understood and its contexts of use, and then consider the pragmatic and hedonic quality of the user experience.

5.1 Breadth in Understanding

As an experimental offering from the service provider, ChatGPT users are encouraged to explore and use ChatGPT in diverse ways. Consequently, it is not surprising that our participants exhibit significant variation in their understanding of ChatGPT.

For all users, the technical functionality of ChatGPT was an important defining characteristic. This indicates current users' awareness of the importance of the technological advances enabling ChatGPT. At the same time, a substantial proportion of participants also understood ChatGPT in terms of its potential uses and purposes and its interaction capabilities. These findings may indicate that current users are part of an ongoing meaning-making process [33], identifying what ChatGPT and, by extension, conversational AI, mean in their lives. ChatGPT appears to be construed as a novel technology that potentially holds a significant role for users through its useful applications and manner of interaction.

5.2 Context of Use – a Need to Add a New Dimension to Conversational AI Taxonomies?

ChatGPT is used in various contexts, including work, education, and daily life. The use cases are diverse, ranging from creative and entertaining activities to goal-oriented tasks like information retrieval, software development and writing support. We find the wide usage to be interesting, in part as this demonstrates that ChatGPT is used for more than a narrow set of use cases such as software development [42].

Our results suggests that ChatGPT goes beyond traditional conversational user interfaces, which typically focus on specific tasks such as providing factual information, controlling home devices, or delivering weather forecasts [34], and social conversational user interfaces that emphasize social interactions [35]. With conversational AI, a distinction between task-oriented and socially oriented conversational user interfaces [15] may potentially become less relevant. ChatGPT is relevant for a broad and multi-faceted set of use-cases, something that may challenge existing taxonomies of conversational user interfaces where such user interfaces are considered more limited in scope, such as the typology proposed by Følstad et al., [41]. As noted by Roller et al., [4] the aim of conversational AI should be to "blend entertaining and knowledge while making others feel heard." The participant feedback on ChatGPT suggests that conversational AI may be on its way toward realizing this vision.

5.3 Pragmatic Attributes: Practical Use for the Win

Good experiences are often practical. Most participants highlighted positive experiences when using ChatGPT for practical purposes. Likewise, negative experiences were often associated with breakdowns in pragmatic quality. ChatGPT appears to enable positive experiences when alleviating tasks users need to get done, such as supporting work or school-related tasks, assisting with writing, or enhancing texts, producing code, and supporting decision-making. This interest and willingness to take up ChatGPT for pragmatic gains contrast OpenAI's descriptions of ChatGPT as a research experiment with many unresolved flaws [36]. From such a description, one might expect that ChatGPT would trigger positive user experiences mainly associated with exploration and play. However, our findings contradict this assumption and suggest that good user experiences typically arise when ChatGPT makes users' lives easier. This finding aligns with previous research on conversational user interfaces which suggests the importance of pragmatic quality [13] and highlights that conversational AI providers need to be mindful of users being inclined to take up this technology for practical purposes before it is ready for such use.

The importance of concrete, detailed, and comprehensive answers. We find that good user experiences are often associated with ChatGPT's ability to provide detailed and specific information. In contrast, negative user experiences are triggered in cases where ChatGPT is incapable of doing so. While a tendency to provide only high-level generic information has been acknowledged as a problem for the user experience of conversational user interfaces [27], existing conversational design guidelines often highlight the importance of presenting content in brief messages to avoid overwhelming the user [17, 27].

Our findings show that this may not apply to conversational AI such as ChatGPT. Rather, an inability to offer detailed information may lead to poor user experience. This preference for comprehensive and detailed responses may be due to ChatGPT being used for relatively complex tasks requiring extensive detail. Furthermore, some participants indicated that ChatGPT presents the conversations in a way that makes it easy to digest longer responses by way of, e.g., clear structure and plain language.

The introduction of ChatGPT may represent a shift where conversational user interfaces extend to in-depth content creation. This will significantly alter our conception of how people consume conversational content, and thus our understanding of conversational design and design guidelines.

Challenging to formulate requests – need for prompt-writing support? Although ChatGPT appears decent at understanding its users, some participants have reported poor experiences when it fails to comprehend their requests. This often led them to rephrase their query, which some found problematic and cumbersome. We find this interesting as one of the main technology companies, Microsoft, has argued that one benefit of services such as ChatGPT is that "the chat experience empowers you to refine your search until you get the complete answer you are looking for by asking for more details, clarity, and ideas." [37]. While this

appears to be considered positive in certain contexts, e.g., for educational purposes, it may be experienced as frustrating when users expect ChatGPT to offer the desired response on the first shot.

To avoid frustrating user experiences due to a need for rephrasing, prompt-writing support may be a relevant area of future research – in terms of intelligent support for prompt-writing and optimizing the process of prompt refinement. For example, by having mechanisms that detect use of inefficient prompts and suggest alternatives based on assumptions regarding user intent. Research on prompt-writing support is already seen for other types of generative AI [38] and will likely be a relevant area of research for conversational AI.

Convincingly wrong. Trust is considered important for user experience of other conversational user interfaces [28, 29], and it is not surprising that we find this to matter with ChatGPT. Our results specifically demonstrate how biased or made-up information negatively influences user experience.

In previous research on conversational user interfaces, such as intent-based chatbots, providing correct information has concerned whether the right intent has been identified, with the corresponding delivery of a predetermined and quality-assured response [12]. For intent-based chatbots, any mismatch between the user request and the output tends to be substantial and easy to detect. With conversational AIs, erroneous or non-factual content may be presented as fact, often in a convincing manner [39]. Uncovering misinformation may require expertise, which the user may not have. Particularly, as lack of expertise could be why users ask ChatGPT in the first place. LLMs have previously caused substantial concern due to their ability to generate high-quality disinformation [40]. Substantial research effort will be needed to understand how users deal, or fail to deal, with erroneous or misleading output from conversational AI, and to develop support for users to avoid such output being generated or posing a problem.

5.4 Hedonic attributes: Not the Main Attraction, but of Importance

Though less prominent, we find that hedonic attributes, such as entertaining interactions, support for creativity, and leaving the user impressed or surprised, contribute to positive user experiences. Failing to support creative activities may cause poor user experiences.

The "wow" factor is prominent – interplay between pragmatic attributes and hedonic experiences. We find it noteworthy that several participants expressed that ChatGPT outperformed their expectations, which was typically reported as surprising or impressive. This finding contrasts with the experiences reported in previous studies of conversational user interfaces, where users have been inclined to report disappointment rather than positive surprise [21]. We did not ask the participants why ChatGPT was associated with such positive sentiment. However, it may be interpreted as users closing the "gulf of expectation" as conversational user interfaces have become more commonplace. The positive sentiment may suggest that ChatGPT performs better than expected.

Furthermore, considering the identified prominence of the pragmatic attributes of the ChatGPT user experience, the positive sentiment associated with ChatGPT may suggest an interesting interplay between pragmatic and hedonic attributes where the latter arises because of the first. That is, leaving the users impressed by its capabilities seems important to the hedonic quality and, by extension, user experience – at least during initial interactions.

Supporting creative activities is important. A potential benefit of conversational AI, such as ChatGPT, is its ability to support various use cases. This includes use cases that may be novel to conversational user interfaces, such as creative activities. Our study shows that users may enjoy ChatGPT for creative interactions. Similarly, studies of companion chatbots have demonstrated that users frequently participate in creative role-playing [22]. Exploration and play are generally beneficial for users to understand the technology and identify potential use cases [21, 22]. Therefore, allowing for exploration, creativity, and play may be helpful for conversational AI to provide value to users, even before goal-oriented use cases are established. Moreso, supporting creative activities may attract users that otherwise would not be interested in the technology. Designing for hedonic attributes is important, even in cases where they may not be the primary focus [13].

5.5 Limitations and Future Research

Although our study is an important step towards gaining needed knowledge of conversational AI user experience, it also has limitations that indicate the need for future research. First, while our questionnaire approach enabled us to target a large group, nuances and details may have been missed. Therefore, we recommend that future studies employ in-depth interviews to investigate user experiences more thoroughly. Second, our study does not address potential cultural differences in conversational AI user experience as our sample was limited in size and geography.

Future studies should aim to include a more diverse sample to gain a more comprehensive understanding of conversational AI user experience across cultures. Third, as our study was conducted while ChatGPT was still in its infancy, future studies are needed to uncover changes as users move from initial to stable use. Finally, as our study focused solely on ChatGPT, it will be important to extend this line of research to other conversational AI likely to emerge. For example, Microsoft is currently applying a GPT-integration in Bing, which supports live data and source references.

5.6 Conclusion

Our results demonstrate that ChatGPT user experience is influenced by pragmatic attributes, such as provision of useful and detailed information to ease work- or school-related tasks. Additionally, user experience is impacted by hedonic attributes, such as entertainment and creative interactions, and interactions leaving users impressed or surprised. Our study further demonstrates that ChatGPT may challenges established assumptions regarding conversational user interfaces. That is, ChatGPT offers seamless switching between domains and types of tasks. Moreover, ChatGPT seems to move conversational interaction towards more detailed and comprehensive content, something that markedly contrasts previous conversational user interfaces.

ACKNOWLEDGMENTS

This work was partially supported by the Research Council of Norway, and the Norwegian Media Authority (Grant number: 00374833)

Note: The authors used ChatGPT to improve sentences and structure in small parts of this paper.

REFERENCES

[1] van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R. and Bockting, C. L. ChatGPT: five priorities for research. Nature, 614, 7947 (2023), 224-226.
[2] Hu, K. ChatGPT sets record for fastest-growing user base - analyst note. 2023.
[3] Adiwardana, D., Luong, M.-T., So, D. R., Hall, J., Fiedel, N., Thoppilan, R., Yang, Z., Kulshreshtha, A., Nemade, G. and Lu, Y. Towards a human-like open-domain chatbot. arXiv preprint arXiv:2001.09977 (2020).
[4] Roller, S., Dinan, E., Goyal, N., Ju, D., Williamson, M., Liu, Y., Xu, J., Ott, M., Shuster, K. and Smith, E. M. Recipes for building an open-domain chatbot. arXiv preprint arXiv:2004.13637 (2020).
[5] Shrivastava, R. ChatGPT is coming to a customer service chatbot near you. (2023).
[6] Mollick, E. ChatGPT is a tipping point for AI. Opinion piece, 2022.
[7] Følstad, A. and Taylor, C. Investigating the user experience of customer service chatbot interaction: a framework for qualitative analysis of chatbot dialogues. Quality and User Experience, 6, 1 (2021), 1-17.
[8] ISO 9241-210:2019. Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems, 2019.
[9] Law, E. L.-C. and Van Schaik, P. Modelling user experience–An agenda for research and practice. Interacting with computers, 22, 5 (2010), 313-322.
[10] Bargas-Avila, J. A. and Hornbæk, K. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. CHI, 2011.
[11] Kocaballi, A. B., Laranjo, L. and Coiera, E. Understanding and measuring user experience in conversational interfaces. Interacting with Computers, 31, 2 (2019), 192-207.
[12] Haugeland, I. K. F., Følstad, A., Taylor, C. and Bjørkli, C. A. Understanding the user experience of customer service chatbots: An experimental study of chatbot interaction design. International Journal of Human-Computer Studies, 161 (2022), 102788.
[13] Følstad, A. and Brandtzaeg, P. B. Users' experiences with chatbots: findings from a questionnaire study. Quality and User Experience, 5, 1 (2020), 3.
[14] Hassenzahl, M. The thing and I: understanding the relationship between user and product. Funology 2: from usability to enjoyment (2018), 301-313.
[15] Chen, H., Liu, X., Yin, D. and Tang, J. A survey on dialogue systems: Recent advances and new frontiers. Acm Sigkdd Explorations Newsletter, 19, 2 (2017), 25-35.
[16] Følstad, A. and Skjuve, M. Chatbots for customer service: user experience and motivation. CUI, 2019.
[17] Oesterreich, T. D., Anton, E., Schuir, J., Brehm, A. and Teuteberg, F. How can I help you? Design principles for task-oriented speech dialog systems in customer service. Information Systems and e-Business Management (2022), 1-43
[18] Mogaji, E., Balakrishnan, J., Nwoba, A. C. and Nguyen, N. P. Emerging-market consumers' interactions with banking chatbots. Telematics and Informatics, 65 (2021), 101711.
[19] Melián-González, S., Gutiérrez-Taño, D. and Bulchand-Gidumal, J. Predicting the intentions to use chatbots for travel and tourism. Current Issues in Tourism, 24, 2 (2021), 192-210.
[20] van der Goot, M. J., Hafkamp, L. and Dankfort, Z. Customer service chatbots: A qualitative interview study into the communication journey of customers. Springer, Conversations, 2020.
[21] Luger, E. and Sellen, A. " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. CHI, 2016.
[22] Skjuve, M., Følstad, A., Fostervold, K. I. and Brandtzaeg, P. B. A longitudinal study of human–chatbot relationships. International Journal of Human-Computer Studies, 168 (2022), 102903.
[23] Brandtzaeg, P. B., Skjuve, M., Dysthe, K. and Følstad, A. When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots. ACM press, CHI, 2021
[24] Ta, V., Griffith, C., Boatfield, C., Wang, X., Civitello, M., Bader, H., DeCero, E. and Loggarakis, A. User Experiences of Social Support From Companion Chatbots in Everyday Contexts: Thematic Analysis. Journal of medical Internet research, 22, 3 (2020), e16235.
[25] Shum, H.-Y., He, X.-d. and Li, D. From Eliza to XiaoIce: challenges and opportunities with social chatbots. Frontiers of Information Technology & Electronic Engineering, 19 (2018), 10-26.
[26] Fitzpatrick, K. K., Darcy, A. and Vierhile, M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR mental health, 4, 2 (2017), e7785.
[27] Chaves, A. P. and Gerosa, M. A. How should my chatbot interact? A survey on social characteristics in human–chatbot interaction design. International Journal of Human–Computer Interaction, 37, 8 (2021), 729-758.
[28] Nordheim, C. B., Følstad, A. and Bjørkli, C. A. An initial model of trust in chatbots for customer service—findings from a questionnaire study. Interacting with Computers, 31, 3 (2019), 317-335.
[29] Skjuve, M., Følstad, A., Fostervold, K. I. and Brandtzaeg, P. B. My Chatbot Companion-a Study of Human-Chatbot Relationships. International Journal of Human-Computer Studies, 149 (2021), 102601.
[30] Go, E. and Sundar, S. S. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. Computers in Human Behavior, 97 (2019), 304-316.
[31] Palan, S. and Schitter, C. Prolific. ac—A subject pool for online experiments. Journal of Behavioral and Experimental Finance, 17 (2018), 22-27.
[32] Clarke, V., Braun, V. and Hayfield, N. Thematic analysis. Qualitative psychology: A practical guide to research methods, 3 (2015), 222-248.
[33] Pinch, T. J. and Bijker, W. E. The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. Social studies of science, 14, 3 (1984), 399-441.
[34] Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., Williams, H., Sorsche, P., Hirsch, D., Li, Q. and Martinez, A. Talk to me: Exploring user interactions with the Amazon Alexa. Journal of Librarianship and Information Science, 51, 4 (2019), 984-997.
[35] Ta-Johnson, V. P., Boatfield, C., Wang, X., DeCero, E., Krupica, I. C., Rasof, S. D., Motzer, A. and Pedryc, W. M. Assessing the Topics and Motivating Factors Behind Human-Social Chatbot Interactions: Thematic Analysis of User Experiences. JMIR Human Factors, 9, 4 (2022), e38876.
[36] AI, O. ChatGPT: Optimizing Language Models for Dialogue. (2023).
[37] Mehdi, Y. Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web. (2023).
[38] Dang, H., Mecke, L., Lehmann, F., Goller, S. and Buschek, D. How to Prompt? Opportunities and Challenges of Zero-and Few-Shot Learning for Human-AI Interaction in Creative Applications of Generative Models. arXiv preprint arXiv: 2209.01390 (2022).
[39] Bender, E. M., Gebru, T., McMillan-Major, A. and Shmitchell, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, March 2021.
[40] Sedova, B. B. A. L. M. M. K. Truth, Lies, and Automation. Center for security and emerging technology, 2021.
[41] Følstad, A., Skjuve, M., & Brandtzaeg, P. B. (2019). Different chatbots for different purposes: towards a typology of chatbots to understand interaction design. In Internet Science: INSCI 2018 International Workshops, St. Petersburg, Russia, October 24–26, 2018, Revised Selected Papers 5 (pp. 145-156). Springer International Publishing.
[42] Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). " I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data. arXiv preprint arXiv:2212.05856.
[43] Hornbæk, K., & Hertzum, M. (2017). Technology acceptance and user experience: A review of the experiential component in HCI. ACM Transactions on Computer-Human Interaction (TOCHI), 24(5), 1-30.