Bridging the gap: user expectations for conversational AI services with consideration of user expertise

Daphne Greiner
Sorbonne School of Management (PRISM Sorbonne), Paris, France, and
Jean-François Lemoine
ESSCA School of Management, Angers, France and Sorbonne School of Management (PRISM Sorbonne), Paris, France

Abstract

Purpose – Past research has emphasised the potential for conversational artificial intelligence (AI) to disrupt services. Conversely, the literature recognises customer expectations as fundamental to service quality and customer satisfaction. However, the understanding of users' expectations for conversational AI services is currently limited. Building upon previous research that has underscored the importance of users' expertise, this study aims to provide valuable insights into the expectations of users with varying levels of expertise.

Design/methodology/approach – Forty-five semi-structured interviews were conducted, on three populations: experts, quasi-experts and non-experts from various countries including Japan, France and the USA. This includes 10 experts and 11 quasi-experts, as in professionals in conversational AI and related domains. And 25 non-experts, as in individuals without professional or advanced academic training in AI.

Findings – Findings suggest that users' expectations depend on their expertise, how much they value human contact and why they are using these services. For instance, the higher the expertise the less anthropomorphism was stated to matter compared to technical characteristics, which could be due to a disenchantment effect. Other results include expectations shared by all users such as a need for more ethics including public interest.

Originality/value – The study provides insights into a key yet relatively unexplored area: it defines three major expectations categories (anthropomorphic, technical and ethical) and the associated expectations of each user groups based on expertise. To the best of the authors' knowledge, it also highlights expectations never detected before as such in the literature such as explainability.

Keywords AI services, Conversational AI, AI marketing, User expectations, Expertise, Anthropomorphism

Paper type Research paper

Introduction

Previous endeavours to implement conversational artificial intelligence (AIs) (e.g. chatbots, virtual agents, etc.) have often fallen short of meeting user expectations (Kocielnik *et al.*, 2019, May; Riveiro and Thill, 2021). The "gulf" between user expectations and experience of conversational AI (Luger and Sellen, 2016, May) has led to notable unsuccessful outcomes (Davis, 2016; Wolf *et al.*, 2017; O'Neal, 2019; Zemčík, 2020). Major players in the industry, such as Google and Microsoft, have experienced these setbacks. In 2022, Google Duplex encountered a shutdown primarily due to criticism on its perceived deceptiveness and lack of ethical consideration (O'Leary, 2019; O'Neal, 2019). Notably, the AI never explicitly disclosed its non-human identity, despite exhibiting behaviour and voice that closely resembled that of a human (op. cit.). Similarly, Microsoft faced controversy in 2016 with "the tragic tale" of Tay Bot (Davis, 2016), as the company failed to anticipate user behaviour, particularly online trolling, which resulted in the AI learning and adopting racist and sexist tendencies within 24 h and consequently being shutdown (Wolf *et al.*, 2017; Zemčík, 2020).

In both instances, the comprehensive examination of user expectations appeared to have been insufficiently considered. Considering user expectations could have potentially averted or lessened the impact of these expensive occurrences; as Wolf *et al.* (2017) argue "we should have foreseen this". Likewise, the incident involving Google Duplex ("is Google Duplex too human?", O'Neal, 2019) could have also been further alleviated by considering the Uncanny Valley phenomenon which explains how machines that are too human-like or eerily so can elicit negative user responses (Mori *et al.*, 2012).

Current research on conversational AI tends to focus on either technical aspects or questions around its nature, research directions, its impact or disruption (Bock *et al.*, 2020). This emphasis can be attributed to the rapid expansion of these services within a relatively short timeframe (op. cit.) and could also be due to a lack of funding for AI research outside engineering. Consequently, there is a scarcity of research that

The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/0887-6045.htm**

Journal of Services Marketing
39/2 (2025) 76–94

[DOI 10.1108/JSM-02-2024-0056]

Received 20 February 2024
Revised 23 July 2024
23 October 2024
Accepted 27 October 2024

comprehensively investigates users' overall expectations for conversational AI, to improve their design and user experience.

Most studies have focused on specific anthropomorphic features like voice and appearance when examining user expectations, often neglecting the impact of users' expertise and intent. This research gap has resulted in conflicting findings regarding the desired levels of anthropomorphism and technical attributes, such as intelligence. While some studies on users' overall expectations reveal a disparity between what users want and what is currently available (Kocielnik *et al.*, 2019, May; Riveiro and Thill, 2021), none of these studies have incorporated the existing body of literature on services. This literature provides a clear definition of AI in the service context and identifies areas for future exploration (Huang and Rust, 2018; Bock *et al.*, 2020). In essence, conversational AIs have been predominantly viewed as technical tools rather than services that require an understanding of users' needs and expectations.

The primary focus of this research is to examine the expectations of users based on their level of expertise and, consequently, their intended use or purpose for conversational AI services. Essentially, we aim to understand what users expect in terms of characteristics and how these expectations are influenced by their level of expertise. In other words: What are users' expectations for conversational AI services? And how does their expertise impact their expectations? For instance, a researcher might have different expectations than a gamer using AI for recreation. While it is known that AI expertise can influence attitudes and behaviours (Moreira, 2022), there is a lack of research on how it affects expectations for conversational AI services, which to the best of our knowledge was never explored.

To answer the questions brought forward, 45 semi-structured interviews were conducted in 2021 on three populations: experts, quasi-experts and non-experts. This methodology, though uncommon, has been used by prior research (Hu *et al.*, 2022) and was recommended by Quivy and Van Campenhoudt (1995). The interviews, conducted internationally via video calls due to COVID-19 restrictions, included 10 experts, 11 quasi-experts (professionals in AI or related fields) and 25 non-experts (individuals without advanced AI training). Detailed methodology and sample information are provided in subsequent sections.

The present study contributes to the existing literature in different fields, including Web atmosphere (Lemoine, 2012), AI marketing and experience (Scown, 1985) and service AI (Bock *et al.*, 2020). It expands upon the existing literature by offering more detailed insights into certain findings including how the desired level of anthropomorphism may vary depending on the individual's level of expertise. In addition, the paper uncovers unexplored areas by identifying two primary types of characteristics that users have expectations on, in addition to ethical considerations, namely, anthropomorphic versus purely technical or functional characteristics such as speed and availability.

On a managerial note, these results matter because they could help direct funds on characteristics that matter most to targeted users. In addition, using these findings could help lower risk of failure by providing a more adequate design and experience of conversational AI services based on common expectations and the targeted audience's level of expertise.

In the first section the study's theoretical grounding is presented, providing insight on what is meant by conversational AI services and expertise, and what was known prior to this study on users' expectations for conversational AI. Secondly, the methodology is described as well as the three populations questioned, and the results, going in depth in what expectations users questioned had expressed. Finally, a discussion follows comparing these results to past studies and most importantly, providing potential explanations on findings with the example of the disenchantment effect.

Literature review

Definition of conversational artificial intelligence services and marketing contributions

The term AI, believed to have been first coined by McCarthy in 1955, covers different realities and can be defined simply as an entity capable of receiving, interpreting and learning from external inputs to simulate one or more forms of human intelligence (O'Regan, 2013; Rajaraman, 2014). What is referred to as conversational AI is a subfield of AI, it encompasses multiple facets such as Natural Language Processing and Automatic Speech Recognition, to converse in a natural language with human users (Visakh *et al.*, 2023, June). Although some conversational AIs have been used to exchange with one another. Since the creation of Eliza Therapist by Joseph Weizenbaum (1966), believed to be the first chatbot, a variety of names have been used for conversational AIs. These names such as "virtual agent", "vocal assistant", both depict a diversity in forms of conversational AI but also a lack of consensus on the terminology (Agarwal *et al.*, 2022). McTear (2020), in a work referenced by more than 200 studies, refers to these technologies as dialogue systems, while also using the terms conversational agents, conversational AI and chatbots. The book offers a comprehensive historical analysis of these technologies and discusses various issues in an integrated manner, highlighting numerous design challenges that are prevalent across all conversational AIs, including gender-specific biases and concerns related to the safeguarding of customer data and privacy.

The service literature does not provide a specific definition of conversational AI services. However, the concept of service AI is clearly defined by Bock *et al.* (2020). According to their definition, service AI refers to the utilisation of technology to deliver value in both internal and external service environments. This is achieved through the ability to adapt flexibly, enabled by capabilities such as sensing, learning, decision-making and taking appropriate actions. Conversational AI services, along with their associated technologies such as chatbots, virtual agents and social companions, have been characterised in academic literature as both a unique service offering (Nadarzynski *et al.*, 2019) and as components of a more extensive array of services. This broader categorisation includes digital or e-services (Chung *et al.*, 2020) and sometimes integrates them into specific industry services, exemplified by the tourism sector (Ukpabi *et al.*, 2019).

Although conversational AI services as a whole have yet to be thoroughly explored in the service context, the marketing

literature has greatly contributed to their understanding. Perhaps research has moved faster than managerial practices in this perspective, as the term AI experience was already coined by S.J. Scown in 1985, and the term AI marketing emerges in the literature as a sub-category of digital marketing with great potential (Haleem *et al.*, 2022). Sometimes perceived as a purely customer support technology, they are also products used to motivate users, help employees and much more. For instance, in health care, conversational AIs have proven to be valuable companions to isolated patients (Fournier-Tombs and McHardy, 2023). In a commercial context, J-F. Lemoine (2012) described them as social links for online stores that can enhance online experience or Web atmosphere. However, this social link depends on the conversational AI in question because while users expect them to behave and communicate like humans (Jenkins *et al.*, 2007), they quickly notice the shortcomings of an AI that is too human but not quite so (Uncanny Valley hypothesis, Mori Masahiro, 1970).

Thus, the concept of anthropomorphism is strongly linked to conversational AI and has been of particular interest to marketing scholars. Guthrie (1993) defines it as the attribution of human characteristics to a non-human phenomenon. This definition, which serves as a point of reference, was enriched by the work of several researchers, notably Duffy (2003), Nowak and Biocca (2003) and Kwan *et al.* (2008). By combining their contributions, the following definition of anthropomorphism emerges: attribution of human or animalistic characteristics (both in appearance and behaviour) to non-human entities to rationalise our surroundings. This implies that anthropomorphism can be attributed by designers of conversational AI but also by their users in the way they perceive them (e.g. this machine is annoying). In the context of AI and robot services, anthropomorphism has been found, amongst other things, to drive trust, intention to use and enjoyment (Van Pinxteren *et al.*, 2019).

The marketing literature on expectations for ethical and technical aspects of conversational AI is scarce. Crossroads are met between these notions and anthropomorphism in conversational AI services, for instance with quoted rules such as not intentionally misleading or manipulating consumers (Mozafari *et al.*, 2022). Unfortunately, how to achieve this whilst maintaining service quality has not been thoroughly explored in the literature as ethics in service AI is one of the research areas in need of further investigation (Bock *et al.*, 2020).

User expertise and expectations for conversational artificial intelligence services

Little is known about conversational AI users, as in what are their attitudes, expectations and intent for using these services, nor their preferences (Bergner *et al.*, 2023). Most studies focus on causal effects of anthropomorphic characteristics and draw conclusions about user expectations from these findings. Such is the case for most studies quoted in this section as this research area is relatively new. Furthermore, the diverse sub-categories within conversational AI, such as "chatbots" and "virtual assistants" (Bock *et al.*, 2020; Seaborn *et al.*, 2021), means researchers are forced to use multiple terms during literature exploration, thereby increasing the complexity of the search process.

Studies which provide clues on user expectations for anthropomorphic traits of conversational AI, generally focus on one characteristic such as voice, appearance or, more rarely, personality (Koike and Loughnan, 2021). Practically speaking, this means they often overlook the significance of combining these features to create coherent and credible entities, such as the Miss Piggy chatbot on Messenger, launched by Disney in 2015 to promote the Muppets show. And more recently, the mobile robot, Miroka, which features strong personality traits, introduced in 2023 by the French start-up Enchanted tools. The value of studying anthropomorphic characteristics as a whole, often through the lens of conversational AI personality, has been depicted by a handful of researchers (Bock *et al.*, 2020; Ait Baha *et al.*, 2023). This can be understood in two ways: conversational AIs can have their own constant personality such as one based on a real or fictional character like the Miss Piggy chatbot, or conversational AIs can adopt adaptive personalities, as in adapt their inputs to users' detected personality. On the first note, Medhi Thies *et al.* (2017) underlined the value of more coherent or well-rounded conversational AIs with clear given personalities and found that young Indian users wanted them to work both as useful tools and as friends or companions. Other studies highlight the need to carefully design the personality of AIs to ensure a consistent set of characteristics (Smestad and Volden, 2019). This also applies in a service context, as depicted by Van Pinxteren *et al.* (2020) with the example of how expectations on appearance may affect behavioural expectations. In the same study, authors underline the need to distinguish users by stating that expectations may vary depending on user groups (e.g. users with high and low affinity for technology).

Notable differences between how AI is approached by expert versus non-expert users have been brought to light in the existing literature although how this affects their expectations is not clear. For example, some studies suggest that non-expert users tend to either over-trust or distrust AI systems (Larasati *et al.*, 2021). Contrary to common assumptions, expertise levels are not solely divided between experts and non-experts, as a third group, often understudied is highlighted in the literature, that of quasi-experts (Pfadenhauer, 2009). Exploratory interviews recommended by Quivy and Van Campenhoudt (1995) distinguish and define these populations. Firstly, experts are defined as professionals directly contributing to the field in question. Secondly, quasi-experts are defined as privileged witnesses, for instance, professionals working in related fields. Thirdly, non-experts are defined as users with no professional or extensive academic knowledge of the field or question under study. Conversational AI studies frequently fail to capture these findings, as they typically focus on non-expert users and do not differentiate between users of varying expertise levels.

The current implementation of anthropomorphism as well as technical and ethical aspects of AI may not meet the expectations of users in the coming years. At the crossroad of these three notions sits numerous studies which highlight some users' longing for AIs that possess enhanced emotional intelligence, allowing them to adjust to users' emotional states, replicate emotions and project a friendlier demeanour (Galitsky and Galitsky, 2021; Svikhnushina *et al.*, 2021). Artificial

emotional intelligence (AEI), often combined with social intelligence, seems to be a promising area which can, for instance, help detect and manage customer emotions in automated customer service (Del Prete, 2021) but also comprises technical and ethical challenges to overcome.

In addition to the studies mentioned above, there have been inquiries carried out on the expectations related to the purely technical functionalities of conversational AIs. For instance, one study investigated user expectations regarding the capability of virtual agents to remember and recall conversation history (Richards and Bransky, 2014). The findings revealed that the AI's capacity to remember and retrieve past conversations enhanced the satisfaction, credibility and trust of the surveyed consumers. Conversely, other studies indicate that a virtual agent that responds too quickly, without allowing for pauses, may be perceived as less human-like and, therefore, less desirable (Gnewuch, 2018).

Overall, technical expectations for conversational AIs are less explored in marketing scholarly research and in the context of services. However, other technical expectations relevant to conversational AIs include seamlessness and efficiency (Greiner and Lemoine, 2024). Seamlessness refers amongst other things to how well the AI is integrated with other services (op. cit.). Efficiency is described as encompassing three sub-categories: convenience (accessibility on interface and time availability), user-friendliness (e.g. questions may be poorly phrased) and response accuracy (op. cit.). Whilst notions such as autonomy and explainability are addressed in terms of risks and technological advancements but not as expectations (Rai, 2020).

Similarly, there is a scarcity of studies addressing ethical expectations and demands of users of AI on notions such as transparency (Kerr *et al.*, 2020), and few if any studies for conversational AI users. Current studies on AI ethics mention concepts such as transparency, privacy and data governance as potential expectations of AI users (Felzmann *et al.*, 2019). In addition, accountability, human agency and oversight, diversity, non-discrimination and fairness are also mentioned as potential expectations of AI users (Kerr *et al.*, 2020). However, these notions are not explored as expectations but are suggested as such amongst other findings.

Kerr *et al.* (2020) advise users to be careful but also mention users need to have more realistic expectations of transparency based on technical limitations, which resonates with Luger and Sellen (2016) study on the "gulf" between users' expectations and their experience of conversational agents. It also aligns with a study of Grimes *et al.* (2021) suggesting organisations should set low expectations so users will be less impacted by negative violations. Consistent with these findings, studies on violations between expectations and user experience in AI tend to show that individuals have higher expectations when told an agent is human as opposed to a machine, which beyond ethical concerns provides incentive for more transparent disclosure of conversational AI's nature to users (op. cit.).

What is known about users' expectations for conversational AI remains relatively limited. Studies tend to adopt micro lenses, focusing on one functionality or one community rather than analysing the overall dynamics at play. This means no study, thus far, had provided data on which functionality matters most to whom and why. For instance, should managers care more about anthropomorphic or purely functional characteristics? And how does user expertise matter when it comes to expectations for conversational AI services (e.g. do professors and students have the same needs and wants for conversational AI)?

The classification of conversational AI users is similarly an area that has not been extensively investigated. Recent research, such as that conducted by Gkinko and Elbanna in 2023, categorises AI chatbot users into four distinct groups: early quitters, pragmatics, progressives and persistents. Nonetheless, this study primarily examined the internal utilisation of a single AI chatbot within an organisational context, specifically focusing on users in their capacity as employees. Consequently, the implications of these findings may not be applicable to broader populations (see Table 1 for literature synthesis).

Methodology

Research design

In the previous sections, we highlighted current findings on user expectations for conversational AI services and pending research questions. Mainly, what are the main users' expectations for conversational AI services? And how does their expertise impact their expectations? The need to identify and understand these concepts in-depth, prompted us to use semi-structured

Table 1 Conversational AI literature review synthesis on users' expectations

Characteristics and ethics	User expectations found in the literature
Anthropomorphic characteristics	Varying expectations on voice, appearance and personality depending on users. Overall, they seem to expect a consistent and coherent set of characteristics (e.g. appearance should match behaviour) and expect an equilibrium of the degree of anthropomorphism
Functional characteristics	• Round-the-clock availability • Fast but not too fast responses • High accuracy of responses to queries • Ability to recall past exchanges and continuously learn • Integration of AEI
Ethics	Ethical anthropomorphism, as in AI must not mislead or manipulate users intentionally

Source: Authors' own work

interviews directly with users. This approach allowed us to maintain a certain level of flexibility while addressing the research problem, drawing upon the research avenues outlined in the existing literature to create our questions. This approach has been recommended for investigating complex issues and it offers a middle ground between pros and cons of structured and unstructured interviews (Young *et al.*, 2018).

More precisely, we used a methodology of semi-structured interviews recommended by Quivy and Van Campenhoudt (1995), which involves studying three distinct populations (experts, quasi-experts and non-experts) to gain a thorough understanding of the concept at hand. Although some marketing studies have adopted a similar methodology by examining three populations (such as Kent *et al.*, 2011), the complexity of this approach means interviews on three populations are not commonly employed. Most studies tend to focus on only one or two populations, resulting in a partial reading of concepts. Consequently, readers are required to combine multiple studies to obtain a comprehensive overview, and certain ideas may not emerge.

These interviews were carried out internationally both for theoretical and practical purposes. From a practical standpoint, the limited number and availability of experts and quasi-experts required an international screening. From a theoretical standpoint, Craig and Douglas (2001) indicated a need to conduct additional international marketing research, in part attributed to advances in technology, to consumer's mobility, to the multiplication of firms present internationally and to markets being integrated worldwide (op. cit.). This shift is observable in conversational AI services as they are frequently implemented on an international scale (e.g. ChatGPT, Xiaoice, Siri, Alexa and Kuki), meaning that users have access to the same AIs, although the domestic demand may vary. On this note, prior studies indicate no significant cultural differences in terms of expectations for these services as common expectations emerge such as efficiency and seamlessness. Finally, it is worth noting conversational AIs tend to be standardised as many are outsourced from major international providers such as IBM's AI solution, Watson and smaller ones such as TalkR.ai.

Prior to the interviews, participants were informed that their responses would remain anonymous and that they were intended for academic purposes. They were also made aware that the study aimed to enhance conversational AI services, although experts were provided with more details than non-experts. In addition, participants gave their consent to the interviews being recorded for analysis purposes.

The interview questions primarily consisted of open-ended inquiries on participants' expectations. It also included questions on demographics (gender, age, CSP and place of residency) and prior experience with AI, considered to be an antecedent of trust and adoption (Blut *et al.*, 2021). Additional questions regarding the interviewee's profile were asked such as his/her inclination towards technology and his/her need for human interaction (op. cit.; Van Pinxteren *et al.*, 2020). Finally, experts and quasi-experts were given additional questions pertaining to their work, research and insights.

Data gathering and sample

All interviews were carried out in English or French in 2021 by the same researcher, using video-call systems to adhere to Covid-related restrictions. The researcher recorded and transcribed each interview manually before proceeding with the analysis. On average, each interview lasted approximately 45 min encompassing responses to the interview guide and subsequent discussions involving additional inquiries, variations in interview duration were observed among participants. Notably, interviews with experts tended to be longer (60 min), likely due to the extensive information they shared and their enthusiasm. In addition, a notable contrast emerged within the non-expert population, distinguishing individuals who were inclined to engage with conversational AIs from those who preferred to avoid them. For the latter group, interviews typically lasted around 20 to 30 min, as they expressed limited interest or discomfort with the topic at hand.

Following the naturalistic approach to cross-cultural interviews (Gubrium and Holstein, 2002), participants were selected based on genuine interest, with no financial incentives. Instead, they were motivated by a desire to contribute to research and explore new perspectives. Challenges such as the insider-outsider issue during data collection were not encountered, potentially due to the cross-cultural identity of the interviewer (Tabane and Bouwer, 2006) and the non-sensitive nature of the topic.

The literature does not provide a specific number of participants required for interviews, but it does emphasise semantic saturation as a reliable indicator (Saunders and Townsend, 2016), which we have used. Hence, the number of interviews varies depending on the population being studied, as saturation was reached more quickly with experts. Furthermore, the sample sizes could not be equal due to the significant difference in population sizes. Due to the challenges and limitations in accessing the first two populations, we used two sampling methodologies: snowballing and purposive sampling. While we specifically sought individuals with certain characteristics (detailed below), some interviewees were recommended by others and may not have been accessible otherwise.

The selection criteria for Population 1 focused on professionals, both within and outside of the research field, who were recognised for their contributions to conversational AI and robotics. This encompassed individuals such as creators and engineers involved in the development of conversational AIs, including the co-founder of a renowned conversational AI and an engineer at Facebook AI Research (FAIR), as well as founders of start-ups specialising in conversational AI and seasoned researchers in the field. Notably, a professor and director of an AI research centre in Tokyo also fell within this population. In contrast, the criteria for Population 2 were less restrictive and included professionals working in related fields or professionals who had implemented or used AIs in their work lines. More specifically, this group comprised professors with expertise in related domains, such as computer science, although not specifically in AI, as well as business professionals who possessed knowledge of conversational AI through firsthand experience. For instance, a Head of Growth (digital marketing) who anticipated the integration of a conversational AI on his company's website. Another example of a quasi-expert questioned was a journalist specialising in new technologies for a reputable journal (see Table 2 for sample overview and

Table 2 Description of populations and final sample

Populations and samples	Description
Population 1: experts	Professionals with extensive experience of conversational AI. Profiles included seasoned researchers, creators and founders of conversational AI
11 interviewees	
Organisations they work for	Amelia, the conversational design institute, SAS institute, women in AI, imperial college, Facebook, etc.
Countries of residency	France, Denmark, the UK, the USA, Japan and Australia
Population 2: quasi-experts	Professionals working in related fields to conversational AI, or which have implemented them. Profiles include engineers in related domains, professionals in digital marketing and doctors using AI for patient care
10 interviewees	
Companies they work for	BPCE, Le monde, novo hospital and Panthéon-Assas, etc.
Countries of residency	France, the USA and South Africa
Population 3: non-experts	Individuals with no professional or advanced academic training in AI of all ages and gender with or without prior encounters with conversational AIs, residing on French territory. Fifty-two percent identified as women, 44% identified as men
25 interviews	
Professional status	Students, private chauffeur, airline pilot, architect, singer, comedian, unemployed, entrepreneurs, etc.
Nationality and ethnic background	Malagasy, Lebanese, Chinese and French, from Vietnam, Algeria, Eastern Europe or Portugal

Source: Authors' own work

Table 1, 2 and 3 in the Appendix for detailed notes on the interviewee's profiles).

Methodology for analysing

In accordance with previous research on conversational AIs (Lemoine and Cherif, 2012), we used a thematic analysis approach that encompassed both horizontal and vertical analyses. Each interview was initially examined individually (vertically) to gain a comprehensive understanding of its content. Subsequently, these individual analyses were compared (horizontally) with other interviews to identify common patterns and themes. Prior to coding, detailed notes were taken after each interview to capture the main ideas expressed by the participants and ensure that no questions were left unanswered.

In the analysis conducted, emphasis was placed on the distinct attributes of participants and their expertise in AI, rather than their representation of specific subgroups defined by cultural or gender factors. This methodology was chosen because a culturally focused analysis would necessitate a separate investigation centred on the impact of cultural variances on pre-established expectations, ideally using a quantitative framework to reduce potential biases and cultural stereotypes. A pertinent example is the research by Donthu and Yoo (1998), which used Hofstede's cultural dimensions at the individual level alongside the SERVQUAL scale's service quality dimensions.

Concerning the process of horizontal analysis, it involved the identification of text excerpts that discussed similar themes and the subsequent organisation of these excerpts into nodes to uncover recurring topics. The explicit text was examined using a semantic approach, with minimal interpretation, as outlined by Ozuem *et al.* (2022). Themes were established by grouping text units and gradually combining or dividing them into categories. To assess the relative significance of each characteristic mentioned, the frequency of their occurrence was calculated. This provided an overview of the data during the initial stages of the analysis and offered insights into semantic saturation. However, these frequencies lack statistical representativeness and are, therefore, not included in the final results.

Findings are divided from the discussion to allow readers to independently interpret the results. Nevertheless, during the conceptualisation process, the existing literature was considered, from which readers can observe the derivation of terminology and major themes.

Table 3 User expectations based on expertise level (experts, quasi-experts and non-experts)

	Main user expectations by level of expertise		
	Experts	**Quasi-experts**	**Non-experts**
Technical characteristics	Explainability	Explainability	*N.A.*
	N.A.	Non-intrusiveness	Non-intrusiveness
	Autonomy	Autonomy	Fall-back solution
	Efficiency	Performance	Accurateness, speed and availability
	Added value	Added value	Ergonomics
Anthropomorphic characteristics	Persona	Uniqueness, equilibrium and coherence	Personalisation
			Original and calm voice
			Coherent personality
Ethics	Transparency	Transparency	Transparency
	Regulations	Public interest	Human solution

Source: Authors' own work

Methodology for presenting findings

Qualitative findings presentation may be as significant as data analysis (Bekker *et al.*, 2018). There is ongoing debate in the literature regarding the most effective presentation techniques for qualitative research, particularly in relation to the use of tables to represent qualitative data and the idea of quantification (Zilber *et al.*, 2019).

Following guidelines from prior research stating, "one size doesn't fit all" (op. cit.), a personalised approach was adopted based on documented approaches. The adopted approach, which may be referred to as anthropological, stems from methodological principles from anthropology and qualitative sociology which focus on narrative by providing a holistic representation of the data. Findings are presented as a long description with headings and sub-headings that organise the text according to the points considered significant due to their theoretical implications. In the present study, headings are divided into two main parts: one on the overview of data; a second on the expectations of each population (experts, quasi-experts and non-experts). Sub-headings are organised as follows for each population: overview of expectations, expectations on technical characteristics, expectations on anthropomorphic characteristics and expectations on ethics. Finally, the methodology for presenting findings in this study could also be described as a thematic approach with the concept of "researcher–storyteller", where findings' presentation relies on the narrative of the main message (Bekker *et al.*, 2018).

Findings

Overview of findings

Main findings suggest that expectations for conversational AI depend in part on users' level of expertise, the reason why they are using the AI (user intent) and how much their value human contact. In addition, two types of characteristics emerged from the analysis: anthropomorphic (voice, appearance and personality) versus technical or functional characteristics (speed, availability, etc.) (see Table 3 depicting the main user expectations by level of expertise by the mentioned three categories). The higher the expertise, the higher the expectations for purely technical or functional characteristics were, and the lower the expectations and need for anthropomorphic characteristics (as expressed during interviews). Inversely, the lower the expertise, the higher the expectations and need for anthropomorphic characteristics, and the more precise expectations were. Finally, the lower the expertise the higher the media and works of fiction seemed to influence users' expectations. Regardless, of the users' level of expertise, interviewees indicated wanting improved or smarter conversational AI with more personalisation: both in terms of the exchange (e.g. answer relevancy) and the AI itself (e.g. coherent with brand image). Ethical considerations were also expectations shared by all populations although this did not translate in the same way depending on the user (e.g. privacy versus ethnic representativity issues).

Expert users' expectations for conversational artificial intelligence services

Overview. Overall, experts demonstrated a strong interest in the functionality and potential advancements of conversational AI, focusing on its explicability, autonomy, efficiency and added value, while also highlighting the importance of ethical considerations such as transparency and regulation in offering these services. Their in-depth understanding of these technologies, akin to being familiar with the inner workings of magic tricks or being behind-the-scenes, induced what could be described as a disenchantment effect, where anthropomorphism described as an artifice was not a key expectation unless they were prepared to engage in the charade. Nonetheless, they deemed anthropomorphism and emotional AI as necessary for a wider audience in order amongst other things, to increase trust, user satisfaction and improve user experience.

Expectations on technical characteristics. Explicability was simply described as how explainable was the process leading to the AI's responses. This concept was closely linked to the autonomy of conversational AIs. Experts seemed to be looking forward to more autonomous AIs, particularly in terms of learning capabilities, although concerns over the risks of unsupervised AI were also raised (Director M). They believed that conversational AIs should possess a broad understanding of the world, including commonly known facts and important news, to provide accurate information and engage in high quality conversations, e.g. "the notion of understanding the context is crucial, and is obvious to a human but not to a machine", Director M.

This resonates with the need for conversational AIs to adapt not only to the conversation context but also to its users for instance through user profiling (op. cit.) and to properly detect user intent and manage ambiguity (e.g. "orange" may refer to the colour or fruit, Founder and CEO B). Professor T describes it as partly philosophical "how should we think about something, recognise a word?". Currently experts perceived conversational AIs as limited "command and control" systems (Leading conversational AI creator, Dr L).

It also echoes with their third expectation on efficiency, which emphasised the accuracy of responses and the AI's ability to fulfil its purpose (amongst others Professor N, Director Y), as well as going above and beyond clients' minimum requirements, what Director M refers to as services "performance obligation". On this note, technical experts mentioned concerns over the limits of machine learning, described by Leading conversational AI creator, Dr L. as a brute force requiring large data sets and not permitting a real understanding from AI, in contrast to symbolic AI (Founder and CEO B). They also referred to other technical limits including on natural language understanding (NLU) and voice recognition (op. cit.). On this note, professor T mentions the challenge of conversational AIs integration in robotics, going from command to acting in the real world, and provides a humorous illustration of the limited understanding of some AIs: "You know an AI can recognise humans have two eyes, but if you ask, "how many eyes do I have in my leg" the AI could say "two eyes"". For him this issue partly derives from a lack of communication between experts "the problem is that various forms of AI such as recognition of speech, image, emotion generation or simulation are developed separately".

All these expectations lead to technical expectations on the AI's seamlessness and added value, as in how it differs from a simple search engine or the services of a human agent whilst

ensuring a smooth user experience. This included multiple factors such as consistency (Founder and CEO H), availability, accessibility, usefulness (Founder and CEO B) and offering the user a pleasant experience. "Above all, consumers must not be bored by these chatbots; they must provide something, they must have added value, whatever that may be", Director M. One major risk being providing "a degraded service to save money" (op. cit.).

Expectations on anthropomorphic characteristics. Experts did not always perceive conversational AIs as truly "conversational" as they believed achieving this would require general AI: "at the beginning, chatbots only took FAQs, that's not conversational. Conversational is when I have a real exchange, and behind that I have real intelligence" (Director Y); meaning they viewed these AIs as useful tools and did not display significant interest in anthropomorphic characteristics. These characteristics were instead considered potentially misleading: "even if the chatbot doesn't look like a human, many will perceive it as such" (Founder and CEO H) and "Siri, for example, prompted a film called Her, which is the height of anthropomorphism [. . .] this anthropomorphism created something, a link with humans, which was a big lie" (Leading conversational AI creator, Dr L). However, experts mentioned that these characteristics along with emotional AI such as empathy and brand representativeness (Director Y) may be expected from non-expert-users to improve user experience (Founder and CEO H) and certain metrics such as client satisfaction (Director Y) trust and social presence: "expressing social behaviour and personality [. . .] I think it's essential, depending on what you want to achieve, but it's essential if you want to create the social presence that ensures that virtual characters are appreciated and accepted by users" (Professor N). In addition, attributing a persona to these AIs may well be a requirement to conversation coherence for Founder and CEO H. He explains designers need to make "this technology more human centric [. . .]. For most companies, the main challenge is when you only have engineers that talk to the client. You need a team to understand that the person talking to your AI is not just an IP address, so understanding the human component is crucial".

Expectations on ethics. In terms of ethics, experts primarily called for increased regulation and explicability to prevent prejudice resulting from errors or algorithm bias and for data privacy (amongst others, concerns raised by Director Y, Professor N, Director M): "AI today is not AI but a toolbox [. . .] The hammer can be used for good or ill, hence the need for regulation too", Leading conversational AI creator, Dr L. A third reason for these regulations would be to mitigate the risk of manipulation: "concerns may be when the machine modifies human behaviour [. . .]. An agent who urges me to do sport every day is fine, but one who urges me to vote for this or that president is not." (Professor N).

This concern also relates to notions of cybersecurity, propaganda and accountability, which Founder and CEO H refers to as "duty to honor the trust of users". While they did not delve into specific regulations, they emphasised the accountability of designers and stressed the need for interdisciplinary exchanges. Experts also highlighted a need for more education to lower expectations of these services or make them more realistic for non-expert-users: "The problem today is that we do everything with AI, we talk about it indiscriminately without always understanding it [. . .] as if it were a supreme being", Leading conversational AI creator, Dr L. On this note, Professor N mentions the winter of AI following a disproportionate enthusiasm for AI in the 1980s: "When I was talking to people in the general public, I realised that they didn't know what AI was. And I realised that there was a real need to make people understand what AI was, to avoid falling back into the kind of ecstasy, the kind of crazy dream that we had in the 80 s". According to him, this confusion between expectations and the current capacities of AI can't solely be attributed to the media but also stems from experts using misleading words such as machine learning and artificial neural networks: "When we talk about "machine learning", people imagine that the machine is learning, but it's not. It's doing statistics to try and calculate the best possible score for a decision, or to group objects by similarity in a one-dimensional space. So we are paying a bit of a price for wanting to put funny words on algorithms, which is confusing people". A final expectation on ethics voiced by experts concerned transparency such as the explicit disclosure of the non-human nature of these AIs (Professor N, Founder and CEO H).

Quasi-expert users' expectations for conversational artificial intelligence services

Overview. Quasi-experts held a somewhat critical stance towards conversational AI services, positioning themselves between regular users and professionals in the field. They made a clear distinction between their expectations as professional users and their expectations as recreational users, such as when engaging with conversational AIs in video games. Their main expectations on technical characteristics concerned non-intrusiveness, performance and added value. Whilst key anthropomorphic expectations were uniqueness and coherence. Ethical expectations sometimes connected to anthropomorphic features such as on vocal and physical diversity. However, main ethical expectations concerned transparency, explainability in instances such as medical diagnosis and public interest.

Expectations on technical characteristics. Their technical expectations revolved around three key aspects: non-intrusiveness, performance and added value. Non-intrusiveness referred to the conversational AI remaining passive until it was specifically requested or prompted by user actions: "for example, a customer who has been browsing for some time without making a purchase may need an incentive", Director V. Essentially, it emphasised the importance of the user initiating the conversation.

On the other hand, performance was closely related to the concept of efficiency, as mentioned by experts. However, while efficiency focused on the technical effectiveness of the conversational AI, performance encompassed broader business considerations, such as how well the AI met expectations. Similarly, to experts, quasi-experts mentioned that most AI services could be technically improved (amongst other CTO P and Senior journalist B), such as improved autonomy, contextual understanding and learning. Their expectations on technical characteristics were sometimes let down, as CTO P says "When I think of AI in an ideal way, it's almost something miraculous, you give it information, and it completely processes

it [. . .] Thinking back, each time I had to talk with a chatbot of a company I was always disappointed, I don't know if that's because I'm a developer but I already know what to expect".

Finally, added value encompassed similar aspects as those mentioned by experts, but with additional practical elements such as enhancing productivity (mentioned amongst others by Professor M). Quasi-experts particularly looked forward to enhanced customer service and improved user experience: "If the chatbot doesn't deliver anything, the user experience will plummet, and that's terrible. It has to provide the expected added value", Consultant L. This should be based on a business context, including industry and product category, with the example of the medical field (mentioned amongst others by Engineer N and Doctor V). They believed the way in which these services were used should define how they are designed: "usage takes precedence", Senior journalist B.

Expectations on anthropomorphic characteristics. The main expectation on anthropomorphic characteristics, quasi-experts expressed, was uniqueness and coherence. They described it as how personalised the conversational AI was and how it compared with other available conversational AI services. For instance, Senior journalist B gives the example of a customer orientated conversational AI in the luxury sector, capable of assisting in the creation of personalised perfumes whist maintaining brand image. Consultant L pointed out that conversational AIs are often standardised, which can limit their suitability for specific missions and sectors, such as banking.

The Uncanny valley phenomenon was also mentioned to explain how a balance between technical characteristics and anthropomorphism should be found (mentioned amongst other by CTO P and engineer N). And a crossroad between ethics and anthropomorphism was mentioned when discussing anthropomorphic representativeness or diversity of conversational AI services. Teacher C and PhD candidate A give the example of organisations often adopting female voices for vocal assistants. Manager A explains "to make chatbots attractive, I think you have to make them more human, give them colours, etc. to make people want to communicate with them".

Finally, emotional AI was mentioned but similarly to experts it was mostly done as a projection of what the general population may want. On this note, Director V indicates that it may be an advantage for companies to have more emotionally controlled services but on the other, this could hinder user experience. Teacher C indicates that emotional AI may be an expectation to "provide comfort in daily tasks such as ordering an Uber or checking the weather" but not advisable in contexts such as legislation.

Expectations on ethics. In terms of ethics, quasi-experts expressed their expectations for socially oriented conversational AI services that serve the public interest. For instance, they envisioned conversational AIs being used for the social reintegration of isolated individuals and populations, highlighting their potential for fostering social cohesion, e.g. "in the Covid period, in EPHAD there's been no exchange, some people have let themselves die of loneliness, so yes, AI could help, produce answers, provoke questions", Teacher C. Director V adds "In my opinion, technology should remain at the service of mankind, and should not increase inequalities". Experts mentioning brain implants as interfaces for conversational AI services highlighted beyond unclear medical consequences, ethical and social implications (op. cit.).

Transparency was also mentioned concerning data collection (amongst others by Teacher C) and the disclosure of the AI's non-human nature and the disappointment it may lead to beyond trust issues: "there is a psychological play with setting high expectations and being disappointed with lower results if you think you're talking to a human, but you end up understanding that it's not the case", CTO P. Yet here again, disclosure may not be sufficient as Engineer A mentions "the more technology there is, the less intelligent we become, because it's addictive and it can be a waste of time if you're absorbed in it, it can be a source of distraction".

Similarly to experts, the need for more education of the general public and explainability of algorithms was mentioned by quasi-experts: "Explainability and transparency are key. One needs to understand how it works and be aware of it, especially on social networks", Senior journalist B. However, the complexity of both tasks was acknowledged, amongst others by engineer N, who also mentioned that context made explainability more crucial in certain occurrences: "Medical chatbots misinterpreting symptoms for example could lead to a person not taking vital precautions such as going to the emergency". Yet Teacher C also highlighted how these services properly used could mitigate human error. In other words, it was expected that these services assist rather than replace human services.

Non-expert users' expectations for conversational artificial intelligence

Overview. Non-expert users seemed to have a longer, more precise list of expectations. They would mention specific characteristics or functions such as speed rather than concepts used by experts and quasi-experts (efficiency and performance). The primary technical expectations expressed by non-expert users encompassed the presence of a human fallback solution, accuracy, speed, availability and ergonomics. Although their expectations on anthropomorphism were more diverse, sometimes conflicting based on individual preferences (e.g. tone of voice), three common expectations were highlighted: personalisation, original and calm voice for voice assistants and a coherent personality. Finally, they expressed two main ethical expectations, one on transparency (AI nature disclosure and data management), and a second on when and with whom these services should be used as some expressed a need for human interaction. When a minimum of these expectations was not met, users expressed this may have negative outcomes for organisations such as lower brand esteem: "tests should be carried out on users and in at least 80% of cases, the user should be satisfied. Anything less would work against the brand and tarnish its image", engineer Kevin. Other expectations expressed included non-intrusiveness and seamlessness.

Expectations on technical characteristics. Users mentioned they wanted a human fallback solution when the conversational AI failed to respond accurately or for certain requests, considered too specific or sensitive (mentioned amongst others by student Ruelle). For instance, student Noemie shared her frustrating and painful experience of trying to cancel memberships via conversational AI services following the death of her father.

Unfortunately, interviewees mentioned most conversational AIs seemed to lack this easily accessible human fallback solution which reflected poorly on brand image. They mentioned that prior information given to the AI should directly be transmitted to the human agents to follow-up quickly and that non-human agents still seemed limited compared to works of fiction: "in the movies it's much more advanced", student Celine.

Accurateness, specifically in terms of response accuracy, was another significant expectation held by non-expert users (expressed amongst others by Celine and Kevin). Similar to the need for fallback solutions, they believed that some conversational AIs had a limited scope of responses or scenarios, which hindered its effectiveness in handling various queries. The reliability of the provided information and the quality of service was also mentioned: "relevance and quality of the relationship and information" (student Mina).

On the same note, they mentioned "speed" (expressed, amongst others by Celine and private chauffeur Marc) and availability as key expectations, which echoes with expert's efficiency expectations. Some related verbatims mentioned by manager Sarah were "getting an immediate response" and "it's easy to find, there is no need to search". Overall, the idea behind speed and availability for users, was to save time, for example Mina mentioned: "I'd be quite open to this kind of technology if, for example, it reduced the time spent on a task".

Furthermore, non-expert users expressed a desire for more ergonomic and user-friendly conversational AIs (amongst others by students Lola and Louis). They felt that conversational AIs could sometimes come across as cold or overly formal which put off some users such as Noemie, who gives the example of a virtual assistant that she often uses and expressed her wish for the assistant to remember her and greet her by name. This desire for personalised greetings aligns with the need for a more personalised experience. Finally, what non-expert users referred to as "ergonomics" was how easy the conversational AI was to use. For example, if the conversational AI replied in natural language (preferred) or through conversation options and how accessible it was on an interface.

Expectations on anthropomorphic characteristics. Three main expectations on anthropomorphic characteristics were identified. These expectations include personalisation, an original and calm voice and a coherent personality. Personalisation refers to the desire for more personalised and precise answers, as well as the need for unique anthropomorphic characteristics. This expectation is linked to the desire for original and calm voices ("Hm a calm, steady, intelligible voice", Kevin) that differ from mainstream options like Siri, which sere considered somewhat generic. On this note Mina expressed: "I'd find it rather pleasant to have the voice of a real person rather than the voice of Google, or a toneless robot. A voice with a human presence".

In terms of personality, non-expert users expressed a desire for a coherent personality that aligns with the brand's image and offers a consistent set of characteristics. For instance, if a brand primarily communicates in a formal manner, such as a bank, its virtual agent should also speak formally to users and have a formal appearance, such as professional clothing (expressed amongst others by Celine): "It will depend on the company. For me, the robot will also represent the company's values, and how it is used to interacting with its customers [. . .] A company that uses humour like Michel & Augustin, well, the virtual agent, the chatbot, the robot, I wouldn't mind it being humorous at all.", Celine. Personality traits aligned with the brand's image could make users more inclined to use the conversational AI services again (Kevin).

Finally, users mentioned expectations on appearance, but these were strongly divided between three categories: animalistic appearances, human appearances (such as photos or avatars) and robotic or no appearance (mentioned amongst other by Kevin and Mina). There was also divided opinions in terms of animating or not the AI's appearance: whilst Kevin mentioned preferring static appearance for instant, Ruelle expressed: "I don't like objects that are too static, it's too cold". Overall appearance should be based on client preferences as Kevin illustrates it: "Let's imagine that the chatbot is someone dressed in a tracksuit. You know it's just a chatbot, but you're thinking wow, the team behind this, I don't want to find out about their work".

Expectations on ethics. In addition to anthropomorphic characteristics, non-expert users also expressed a significant ethical expectation, similar to the two first populations: transparency in terms of the AI's nature and data management. A question on what topics can be handled by non-human service providers was also raised with the example of Noemie's grieving experience and with whom as some users may be less inclined to AI, as Celine explained: "I didn't have the answer to my question because [. . .] Um [. . .] I needed human contact" and video-editor Mike, "I prefer to speak with a human" whilst Kevin expresses "no preference, if the chatbot is as sophisticated as the response from a human". Marc adds another dimension to this issue by stating: "with a human you can negotiate with a chatbot you cannot".

No differences in terms of user expectations were attributed to cultural differences. Instead, cultural differences observed due to the international nature of the study concerned adoption and perception which can impact user expectation but with an unclear link, similarly to how motivations may influence user expectations (Daphne and Jean-François, 2024).

Discussion and implications

Prior conversational AI fails have highlighted a disconnection between user expectations and how conversational AI are designed (O'Neal, 2019; Zemčík, 2020). Marketing research could lead to improved managerial practices in conversational AI by providing recommendations and indicating how marketing may be valuable in the creation of future AI services. Currently, leading companies in AI tend to mostly recruit engineers, which can lead to negative outcomes in terms of how the AI is received by its intended users and ethical scandals that go beyond algorithm biases (e.g. vulnerable users such as children or consumers with addictions are not considered). The present study strives to shade some light on user expectations for conversational AI whilst paving the way for future research on other aspects of users of conversational AI (who are they? What do they want or need? Why does what they think matter? And how do they use conversational AI services?).

While chatbots and virtual agents have traditionally been examined within specific contexts, such as customer service or

personal assistance in B2C settings, much of the prior research has focused on the technical and anthropomorphic characteristics of these systems and their effects on users (e.g. Soderlund *et al.*, 2021). In contrast, our paper takes a broader, more integrative approach by examining conversational AI services as a comprehensive domain within AI services (Bock *et al.*, 2020). We expand beyond isolated use cases to consider these systems as embedded components of both everyday life and professional environments, providing a more holistic view of their real-world applications.

Furthermore, this study departs from the traditional focus on how a single stimulus impacts users, shifting towards a deeper exploration of user identities and expectations. This allows us to highlight where design efforts should be concentrated within three key dimensions: technical, anthropomorphic and ethical. By incorporating these novel perspectives, the paper offers a more robust theoretical framework for understanding the diverse needs of users interacting with conversational AIs (including amongst others, chatbots, virtual companions, virtual agents and conversational robots). In other words, instead of starting with a specific stimulus and examining its effect on user expectations, we reverse the approach by first identifying user expectations and then determining how design and stimuli should align to meet those needs.

On another note, the suggested user taxonomy grounded in expertise enhances the current body of literature by providing a versatile framework applicable in diverse contexts, in contrast to earlier research (e.g. Gkinko and Elbanna, 2023), which predominantly concentrated on classifying users as employees within a particular organisation. Furthermore, it posits that existing assumptions about cultural disparities in the realm of AI may be overstated, shifting the focus from cultural disparities to potentially more prevalent parameters such as user intent and expertise.

Main theoretical contributions

From a theoretical point of view, our research extends previous work in different fields of AI and marketing including AI experience (Scown, 1985), Web atmosphere (Lemoine, 2012) and service AI (Bock *et al.*, 2020). Prior studies in marketing have portrayed conversational AIs as useful marketing tools (Kurachi *et al.*, 2018; Cheng and Jiang, 2021) and service enablers (Bock *et al.*, 2020) which resonates with some of our findings such as how these services can be used for different purposes, both personal and professional. And although limited some studies have taken an interest in user expectations. Some of these studies highlight that experience with conversational AI seem to shape customers' expectations and, thus, may play a key role in conceptualising AI services of quality (Tran *et al.*, 2021).

Yet no study, thus far, had unravelled how the expertise of conversational users could influence their expectations. In addition, prior findings suggest that anthropomorphism is expected, but the extent to which a conversational AI should appear humanlike remains unclear (Chandra *et al.*, 2022). The present study provides insight on these questions by indicating that the level of anthropomorphism wanted may depend on the level of expertise of intended users. Interviewees responses tend to show that the higher their expertise, the less anthropomorphism is needed, whereas the non-expert users indicate the opposite: the lower the expertise, the higher the need for anthropomorphism. Furthermore, results indicate that other expectations may also be influenced by users' expertise level including ethical considerations and purely technical characteristics such as speed (time taken to respond). To the best of our knowledge, no study, thus far, had provided an overview of user expectations for conversational AI: anthropomorphic expectations, purely technical expectations and expectations on ethics.

Recent studies on human–machine communication suggest that a conversational AI's effectiveness may depend on the communication context (Tsai *et al.*, 2021). Our findings extend these results by suggesting that expectations for conversational AI seem to depend on the reason why the user is using the AI (user intent). A notable example was in case of tragic events, such as death from a family member, users trying to cancel memberships seemed more inclined to exchange with a human agent than a virtual agent, and their expectations for the virtual agent if no other solution was provided first-hand, were not the same as the ones in more common or pleasant situations. In addition, prior research had found that the user's human interaction need may affect adoption intentions (Song *et al.*, 2022), which aligns with our findings indicating that how much users value human contact may affect their expectations and willingness to use conversational AI services.

Beyond insights on user expectations, the study highlights an interesting phenomenon indicating a possible disenchantment effect for expert users who know AI's backstage and are less willing to play along when it comes to a non-recreational use. This extends the literature on the disenchanted, enchantment theory in technology adoption, in the face of sceptical consumers (Belk *et al.*, 2020).

To summarise, this research represents a pioneering effort in exploring user expectations regarding conversational AI services as the first on this topic. It breaks new ground by examining the varying levels of user expertise in this domain and categorising expectations into anthropomorphic, technical and ethical dimensions, as defined in the existing literature. Notably, it introduces a disenchantment effect to elucidate the variations observed in anthropomorphism and intelligence expectations. Beyond antecedents of trust and adoption motivations, it strives to provide the groundwork for additional research on user expectations in AI services and suggests notions, such as explainability (Rai, 2020) and autonomy, studied in other contexts, are also user expectations. Finally, it provides substantial methodological elements for international studies on AI, which may facilitate future research.

Managerial and societal implications

From a managerial point of view, the study has clear implications for designing conversational AIs that meet user expectations. Mainly that conversational AI design should consider the expectations of their specific target audience and how the AI services are going to be used (defining clear use cases). More precisely, brands creating or trying to implement conversational AIs should have a clear agenda based on their intended users and the main purpose of these AIs. For example, if the AI is meant to provide a personalised shopping experience, the accurateness of its responses and ergonomic considerations may be more important than its speed. These

findings could help managers prioritise the conversational AI's characteristics based on their budget and anticipate how user expectations may evolve in line with AI advancements. In addition, ethical guidelines that users expect from conversational AI services could help building user trust and avoid some pitfalls or even future scandals similar to the Google Duplex case. Finally, taking in consideration, these findings could also improve users' experience with AI and possibly improve their adoption. From conversational AIs in health care, to retail and many other sectors, addressing user expectations for these technologies could improve an organisation's overall service delivery and may positively reflect on its brand image (Bergner *et al.*, 2023).

From a societal perspective, these findings call for public policies to take the first step in regulating various ethical aspects of conversational AIs that organisations have so far overlooked. This includes providing feasible legal frameworks for transparency on data collection and usage within conversational AIs, as well as mandatory divulgation of its nature to its users. It also includes protecting vulnerable users, such as minors, and addressing the sensitive topic of AI ownership, similar to the question of whether Google should be considered a public service? Going further, it could also trigger a debate and plan of actions for AI for public interest, as it was originally intended, and not solely for commercial use. This approach could help ensure that AI development is not only technologically advanced but also aligned with societal values and ethical standards.

Limitations and future research

Nonetheless, the study has a few limitations. Firstly, as it strictly focuses on expectations, it does not provide an in-depth understanding of related notions such adoption motivations and antecedents of trust (Daphne and Jean-François, 2024). However, this could provide ground for future research, as research questions and potential relations between concepts are highlighted in our findings, including how adoption may depend on the user's need for human interaction. Secondly, it can be argued that results on the expectations of each population could be extended which is not possible through a single study. This as well could be a ground for future research focusing, for instance, on one population or one of the expectations highlighted. For instance, how can conversational AI services be more transparent, and what are the possible effects on users? In addition, since very few individuals qualify as experts in conversational AI, higher expertise correlates with fewer interviewees, resulting in an unequal number of interviewees across the compared populations. For future research, this means that doing quantitative studies solely on non-expert users seems more feasible than testing all three populations, this could lead to the creation of persona. Finally, the significance of user proficiency is frequently emphasised in research on AI and other technological advancements (Chinchanachokchai *et al.*, 2021; Inkpen *et al.*, 2023). Nevertheless, there is a notable lack of comprehensive methodology guidelines pertaining to this matter. This paper aims to address this gap by providing valuable insights for future investigations while acknowledging its own limitations in comparison to more widely adopted research methodologies. Finally, future studies could focus on other approaches towards user expectations for conversational AI such as its implications and effects on interactions and trust and user behaviour towards these technologies.

References

Ait Baha, T., El Hajji, M., EsSaady, Y. and Fadili, H. (2023), "The power of personalization: a systematic review of personality-adaptive chatbots", *SN Computer Science*, Vol. 4 No. 5, p. 661.

Agarwal, S., Agarwal, B. and Gupta, R. (2022), "Chatbots and virtual assistants: a bibliometric analysis", *Library Hi Tech*, Vol. 40 No. 4, pp. 1013-1030, doi: 10.1108/lht-09-2021-0330.

Bekker, S. and Clark, A.M. (2018), "Improving qualitative research findings presentations: insights from genre theory", *International Journal of Qualitative Methods*, Vol. 17 No. 1, doi: 10.1177/1609406918786335.

Belk, R.W., Weijo, H. and Kozinets, R.V. (2020), "Enchantment and perpetual desire: theorizing disenchanted enchantment and technology adoption", *Marketing Theory*, Vol. 21 No. 1, pp. 25-52, doi: 10.1177/1470593120961461.

Bergner, A., Hildebrand, C. and Häubl, G. (2023), "Machine talk: how verbal embodiment in conversational AI shapes consumer–brand relationships", *Journal of Consumer Research*, Vol. 50 No. 4, pp. 742-764, doi: 10.1093/jcr/ucad014.

Blut, M., Wang, C., Wünderlich, N.V., *et al.* (2021), "Understanding anthropomorphism in service provision: a meta-analysis of physical robots, chatbots, and other AI", *J. of the Acad. Mark. Sci*, Vol. 49 No. 4, pp. 632-658, doi: 10.1007/s11747-020-00762-y.

Bock, D.E., Wolter, J.S. and Ferrell, O.C. (2020), "Artificial intelligence: disrupting what we know about services", *Journal of Services Marketing*, Vol. 34 No. 3, pp. 317-334, doi: 10.1108/jsm-01-2019-0047.

Chandra, S., Shirish, A. and Srivastava, S.C. (2022), "To be or not to be . . . human? Theorizing the role of human-like competencies in conversational artificial intelligence agents", *Journal of Management Information Systems*, Vol. 39 No. 4, pp. 969-1005, doi: 10.1080/07421222.2022.2127441.

Cheng, Y. and Jiang, H. (2021), "Customer–brand relationship in the era of artificial intelligence: understanding the role of chatbot marketing efforts", *Journal of Product & Brand Management*, Vol. 31 No. 2, pp. 252-264, doi: 10.1108/jpbm-05-2020-2907.

Chinchanachokchai, S., Thontirawong, P. and Chinchanachokchai, P. (2021), "A tale of two recommender systems: the moderating role of consumer expertise on artificial intelligence based product recommendations", *Journal of Retailing and Consumer Services*, Vol. 61, p. 102528, doi: 10.1016/j.jretconser.2021.102528.

Chung, M., Ko, E., Joung, H. and Kim, S.J. (2020), "Chatbot e-service and customer satisfaction regarding luxury brands", *Journal of Business Research*, Vol. 117, pp. 587-595, doi: 10.1016/j.jbusres.2018.10.004.

Davis, E. (2016), "AI amusements: the tragic tale of Tay the chatbot", *AI Matters*, Vol. 2 No. 4, pp. 20-24, doi: 10.1145/3008665.300867.

Del Prete, M. (2021), Emotional artificial intelligence: detecting and managing customer emotions in automated customer service.

Donthu, N. and Yoo, B. (1998), "Cultural influences on service quality expectations", *Journal of Service Research*, Vol. 1 No. 2, pp. 178-186.

Duffy, B. (2003), "Anthropomorphism and the social robot", *Robotics and Autonomous Systems*, Vol. 42 Nos 3/4, pp. 177-190, doi: 10.1016/s0921-8890(02)00374-3.

Felzmann, H., Villaronga, E.F., Lutz, C. and Tamò-Larrieux, A. (2019), "Transparency you can trust: transparency requirements for artificial intelligence between legal norms and contextual concerns", *Big Data & Society*, Vol. 6 No. 1, doi: 10.1177/2053951719860542.

Fournier-Tombs, É. and McHardy, J. (2023), "A medical ethics framework for conversational artificial intelligence", *Journal of Medical Internet Research*, Vol. 25, doi: 10.2196/43068.

Galitsky, B. and Galitsky, B. (2021), "Adjusting chatbot conversation to user personality and mood", *Artificial Intelligence for Customer Relationship Management: Solving Customer Problems*, Springer, pp. 93-127.

Gkinko, L. and Elbanna, A. (2023), "The appropriation of conversational AI in the workplace: a taxonomy of AI chatbot users", *International Journal of Information Management*, Vol. 69, p. 102568, doi: 10.1016/j.ijinfomgt.2022.102568.

Gnewuch, U. (2018), "Faster is not always better: understanding the effect of dynamic response delays in human-chatbot interaction", *Published in European Conference on Information Systems*, available at: www.semanticscholar.org/paper/Faster-is-Not-Always-Better%3A-Understanding-the-of-Gnewuch-Morana/22cbf658ea99b2901b3f6f649e21ef8a3c7a590d

Greiner, D. and Lemoine, J.F. (2024), "Understanding chatbot end-users to improve their design: antecedents of trust, adoption motivations, and expectations", *International Conference on Human-Computer Interaction*, Springer Nature, *Switzerland, Cham*, pp. 194-204.

Grimes, G.M., Schuetzler, R.M. and Giboney, J.S. (2021), "Mental models and expectation violations in conversational AI interactions", *Decision Support Systems*, Vol. 144, p. 113515.

Gubrium, J.F. and Holstein, J.A. (2002), "Handbook of interview research: context and method", Sage publications, available at: https://books.google.fr/books?hl=en&lr=&id=uQMUMQJZU4gC&oi=fnd&pg=PA335&dq=culture+collect+data+interview&ots=V094aCYW3G&sig=QnViU7FmuDZwsAHenPDmAxbYEAk&redir_esc=y#v=onepage&q=culture%20collect%20data%20interview&f=false

Guthrie, S. (1993), "Faces in the clouds: a new theory of religion", available at: https://philarchive.org/rec/GUTFIT

Haleem, A., Javaid, M., Qadri, M.A., Singh, R.P. and Suman, R. (2022), "Artificial intelligence (AI) applications for marketing: a literature-based study", *International Journal of Intelligent Networks*, Vol. 3, pp. 119-132.

Hu, L., Boggess, M. and Shepley, M.M. (2022), "Comparing expert, quasi-expert, and novice evaluations of award-winning design products using the consensual assessment technique", *Creativity Research Journal*, Vol. 35 No. 4, pp. 730-748, doi: 10.1080/10400419.2022.2157591.

Huang, M.H. and Rust, R.T. (2018), "Artificial intelligence in service", *Journal of Service Research*, Vol. 21 No. 2, pp. 155-172, doi: 10.1177/1094670517752459.

Inkpen, K., *et al.* (2023), "Advancing Human-AI complementarity: the impact of user expertise and algorithmic tuning on joint decision making", *ACM Transactions on Computer-Human Interaction*, Vol. 30 No. 5, pp. 1-29, doi: 10.1145/3534561.

Jenkins, M.-C., *et al.* (2007), "Analysis of user interaction with service oriented chatbot systems", *Lecture Notes in Computer Science*, pp. 76-83, doi: 10.1007/978-3-540-73110-8_9.

Kent, M.P., Dubois, L. and Wanless, A. (2011), "Food marketing on children's television in two different policy environments", *International Journal of Pediatric Obesity*, Vol. 6 Nos 2/2, doi: 10.3109/17477166.2010.526222.

Kerr, A., Barry, M. and Kelleher, J.D. (2020), "Expectations of artificial intelligence and the performativity of ethics: implications for communication governance", *Big Data & Society*, Vol. 7 No. 1, doi: 10.1177/2053951720915939.

Kocielnik, R., Amershi, S. and Bennett, P. (2019), "Will you accept an imperfect AI?", *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, doi: 10.1145/3290605.3300641.

Koike, M. and Loughnan, S. (2021), "Virtual relationships: anthropomorphism in the digital age", *Social and Personality Psychology Compass*, Vol. 15 No. 6, p. e12603.

Kurachi, Y., Narukawa, S. and Hara, H. (2018), "AI chatbot to realize sophistication of customer contact points", on ResearchGate [Preprint], available at: www.researchgate.net/publication/327101103_AI_chatbot_to_realize_sophistication_of_customer_contact_points

Kwan, V.S.Y., Gosling, S.D. and John, O.P. (2008), "Anthropomorphism as a special case of social perception: a cross–species social relations model analysis of humans and dogs", *Social Cognition*, Vol. 26 No. 2, pp. 129-142, doi: 10.1521/soco.2008.26.2.129.

Larasati, R., De Liddo, A. and Motta, E. (2021), "AI healthcare system interface: explanation design for nonexpert user trust", *ACMIUI-WS 2021: Joint Proceedings of the ACM IUI 2021 Workshops*, April, Vol. 2903, *CEUR Workshop Proceedings*.

Lemoine, J.-F. (2012), "Pour une présentation du concept d'atmosphère des sites web et de ses effets sur le comportement des internautes", *Marché et Organisations*, Vol. N° 15 No. 1, pp. 169-180, doi: 10.3917/maorg.015.0169.

Luger, E. and Sellen, A. (2016), "'Like having a really bad PA': the Gulf between user expectation and experience of conversational agents", Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp. 5286-5297, doi: 10.1145/2858036.285828.

McTear, M. (2020), "Conversational AI: dialogue systems, conversational agents, and chatbots", *Synthesis Lectures on Human Language Technologies*, Vol. 15 No. 3, pp.1-151, doi: 10.2200/S01131ED1V01Y202204HLT051.

Moreira, A.I.A. (2022), "Expert and non-expert perspectives on AI's impact on skills and competencies and education", Safety and Security Sciences Review, available at: https://biztonsagtudomanyi.szemle.uni-obuda.hu/index.php/home/article/view/258

Mori, M., MacDorman, K.F. and Kageki, N. (2012), "The uncanny valley [from the field]", *IEEE Robotics &*

Automation Magazine, Vol. 19 No. 2, pp. 98-100, doi: 10.1109/mra.2012.2192811.

Mozafari, N., Weiger, W.H. and Hammerschmidt, M. (2022), "Trust me, I'm a bot – repercussions of chatbot disclosure in different service frontline settings", *Journal of Service Management*, Vol. 33 No. 2, pp. 221-245, doi: 10.1108/josm-10-2020-0380.

Nadarzynski, T., Miles, O., Cowie, A. and Ridge, D. (2019), "Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: a mixed-methods study", *Digital Health*, Vol. 5, p. 2055207619871808, doi: 10.1177/2055207619871808.

Nowak, K.L. and Biocca, F. (2003), "The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments", *Presence: Teleoperators and Virtual Environments*, Vol. 12 No. 5, pp. 481-494, doi: 10.1162/105474603322761289.

O'Neal, A.L. (2019), "Is google duplex too human?: Exploring user perceptions of opaque conversational agents", University of Texas Libraries, available at: https://repositories.lib.utexas.edu/items/71ac66eb-be57-4d55-8ed7-2be72f79da0c

O'Regan, G. (2013), "Giants of computing", *A Compendium of Select, Pivotal Pioneers*, *London*, doi:, doi: 10.1007/978-1-4471-5340-5.

O'Leary, D.E. (2019), "Google's duplex: pretending to be human", *Intelligent Systems in Accounting, Finance & Management*, Vol. 26 No. 2, pp. 106-106, doi: 10.1002/isaf.1443.

Ozuem, W., Willis, M. and Howell, K. (2022), "Thematic analysis without paradox: sensemaking and context", *Qualitative Market Research: An International Journal*, Vol. 25 No. 1, pp. 143-157, doi: 10.1108/QMR-07-2021-0092.

Pfadenhauer, M. (2009), At eye level: the expert interview – a talk between expert and quasi-expert, *Interviewing Experts*, Palgrave Macmillan UK, London, pp. 81-97.

Quivy, R. and Van Campenhoudt, L. (1995), "Manuel de recherches en sciences sociales", Dunod. Book, available at: www.eyrolles.com/Entreprise/Livre/manuel-de-recherches-en-sciences-sociales-9782100026562/

Rai, A. (2020), "Explainable AI: from black box to glass box", *Journal of the Academy of Marketing Science*, Vol. 48 No. 1, pp. 137-141, doi: 10.1007/s11747-019-00710-5.

Rajaraman, V.J. (2014), "Father of artificial intelligence", *Reason*, Vol. 19, pp. 198-207, doi: 10.1007/s12045-014-0027-9.

Richards, D. and Bransky, K. (2014), "ForgetMeNot: what and how users expect intelligent virtual agents to recall and forget personal conversational content", *International Journal of Human-Computer Studies*, Vol. 72 No. 5, pp. 460-476, doi: 10.1016/j.ijhcs.2014.01.005.

Riveiro, M. and Thill, S. (2021), "That's (not) the output I expected!" on the role of end user expectations in creating explanations of AI systems", *Artificial Intelligence*, Vol. 298, p. 103507, doi: 10.1016/j.artint.2021.103507.

Saunders, M.N.K. and Townsend, K. (2016), "Reporting and justifying the number of interview participants in organization and workplace research", *British Journal of Management*, Vol. 27 No. 4, pp. 836-852, doi: 10.1111/1467-8551.12182.

Scown, S.J. (1985), "The artificial intelligence experience: an introduction", in Digital Equipment Corporation, available as a Google book at, available at: https://books.google.fr/books/about/The_Artificial_Intelligence_Experience.html?id=BP-EAAAAIAAJ&redir_esc=y

Seaborn, K., *et al.* (2021), "Voice in human–agent interaction", *ACM Computing Surveys*, Vol. 54 No. 4, pp. 1-43, doi: 10.1145/3386867.

Smestad, T.L. and Volden, F. (2019), "Chatbot personalities matters: improving the user experience of chatbot interfaces", Internet Science: INSCI 2018 International Workshops, St. Petersburg, Russia, October 24–26, 2018, Revised Selected Papers Vol. 5, Springer International Publishing, pp. 170-181, doi: 10.1007/978-3-030-17705-8_15.

Soderlund, M., Oikarinen, E.L. and Tan, T.M. (2021), "The happy virtual agent and its impact on the human customer in the service encounter", *Journal of Retailing and Consumer Services*, Vol. 59, p. 102401, doi: 10.1016/j.jretconser.2020.102401.

Song, M., Xing, X., Duan, Y., Cohen, J. and Mou, J. (2022), "Will artificial intelligence replace human customer service? The impact of communication quality and privacy risks on adoption intention", *Journal of Retailing and Consumer Services*, Vol. 66, p. 102900.

Svikhnushina, E., Placinta, A. and Pu, P. (2021), "User expectations of conversational chatbots based on online reviews", *DIS '21: Proceedings of the 2021 ACM Designing Interactive Systems Conference*, 10.1145/3461778.3462125.

Tabane, R. and Bouwer, C. (2006), "The influence of cross-cultural interviewing on the generation of data", *South African Journal of Education*, Vol. 26 No. 4, pp. 553-564, available at: https://hdl.handle.net/10520/EJC32100

Thies, I.M., *et al.* (2017), "How do you want your chatbot? An exploratory wizard-of-Oz study with young, urban Indians", *in Lecture Notes in Computer Science*, pp. 441-459, doi: 10.1007/978-3-319-67744-6_28.

Tran, A.D., Pallant, J. and Johnson, L.W. (2021), "Exploring the impact of chatbots on consumer sentiment and expectations in retail", *Journal of Retailing and Consumer Services*, Vol. 63, doi: 10.1016/j.jretconser.2021.102718.

Tsai, W.S., *et al.* (2021), "Human versus chatbot: understanding the role of emotion in health marketing communication for vaccines", *Psychology & Marketing*, Vol. 38 No. 12, pp. 2377-2392, doi: 10.1002/mar.21556.

Ukpabi, D.C., Aslam, B. and Karjaluoto, H. (2019), "Chatbot adoption in tourism services: a conceptual exploration", *Robots, Artificial Intelligence, and Service Automation in Travel, Tourism and Hospitality*, Emerald Publishing, pp. 105-121, doi: 10.1108/978-1-78756-687-320191006.

Van Pinxteren, M.M.E. *et al.* (2019), "Trust in humanoid robots: implications for services marketing", *Journal of Services Marketing*, Vol. 33 No. 4, pp. 507-518, doi: 10.1108/jsm-01-2018-0045.

Van Pinxteren, M.M.E., Pluymaekers, M. and Lemmink, J.G.A.M. (2020), "Human-like communication in conversational agents: a literature review and research agenda", *Journal of Service Management*, Vol. 31 No. 2, pp. 203-225, doi: 10.1108/josm-06-2019-0175.

Visakh, P.R., Meena, P.N. and Anoop, V.S. (2023), "Conversational artificial intelligence in digital healthcare: a bibliometric analysis", *Multi-Disciplinary Trends in Artificial Intelligence, MIWAI 2023, Lecture Notes in Computer Science*,

Springer, Cham, Vol. 14078, pp. 723-734, doi: 10.1007/978-3-031-36402-0_67.

Weizenbaum, J. (1966), "ELIZA – a computer program for the study of natural language communication between man and machine", *Communications of the ACM*, Vol. 9 No. 1, pp. 36-45, doi: 10.1145/365153.365168.

Wolf, M.J., Miller, K. and Grodzinsky, F.S. (2017), "Why we should have seen that coming: comments on Microsoft's Tay" experiment" and wider implications", *Computers & Society*, Vol. 47 No. 3, pp. 54-64, doi: 10.1145/3144592.3144598.

Young, J., *et al.* (2018), "A methodological guide to using and reporting on interviews in conservation science research", *Methods in Ecology and Evolution*, Vol. 9 No. 1, pp. 10-19, doi: 10.1111/2041-210x.12828.

Zemčík, T. (2020), "Failure of chatbot Tay was evil, ugliness and uselessness in its nature or do we judge it through cognitive shortcuts and biases?", *AI & Society*, Vol. 36 No. 1, pp. 361-367, doi: 10.1007/s00146-020-01053-4.

Zilber, T.B., Amis, J.M. and Mair, J. (2019), "The production of managerial knowledge and organizational theory: new approaches to writing, producing and consuming theory", *Research in the Sociology of Organizations, Emerald Publishing Limited*, Vol. 59, doi: 10.1108/S0733-558X20190000059016.

Further reading

Amendola, A. (2021), "Emotional artificial intelligence: detecting and managing customer emotions in automated customer service", available at: http://elea.unisa.it/handle/10556/6491 (accessed 12 December 2023).

Baby, A. and Kannammal, A. (2020), "Network path analysis for developing an enhanced TAM model: a user-centric e-learning perspective", *Computers in Human Behavior*, Vol. 107, p. 106081, doi: 10.1016/j.chb.2019.07.024.

Beneteau, E., *et al.* (2020), "Parenting with Alexa: exploring the introduction of smart speakers on family dynamics", *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, doi: 10.1145/3313831.3376344.

Brandtzæg, P.B. and Følstad, A. (2017), "Why people use chatbots", *Lecture Notes in Computer Science*, pp, pp. 377–392, doi: 10.1007/978-3-319-70284-1_30.

Cafaro, A., *et al.* (2012), "First impressions: users' judgments of virtual agents' personality and interpersonal attitude in first encounters", *Lecture Notes in Computer Science*, pp. 67-80, doi: 10.1007/978-3-642-33197-8_7.

Chérif, E. and Lemoine, J.-F. (2018), "Les conseillers virtuels anthropomorphes et les réactions des internautes: une expérimentation portant sur la voix du conseiller", *Recherche Et Applications En Marketing (French Edition)*, Vol. 34 No. 1, pp. 29-49, doi: 10.1177/0767370118775963.

Chérif, E. and Lemoine, J.-F. (2019), "Anthropomorphic virtual assistants and the reactions of Internet users: an experiment on the assistant's voice", *Recherche et Applications en Marketing (English Edition)*, , Vol. 34 No. 1, pp. 29-49, doi: 10.1177/0767370118775963.

Dean, A. (2004), "Rethinking customer expectations of service quality: are call centers different?", *Journal of Services Marketing*, Vol. 18 No. 1, pp. 60-78, doi: 10.1108/08876040410520717.

Delecroix, F., Morge, M. and Routier, J.-C. (2012), "A virtual selling agent which is proactive and adaptive", in *Advances in Intelligent and Soft Computing*, pp. 57-66, doi: 10.1007/978-3-642-28786-2_6.

Esposito, A., *et al.* (2019), "Elder user's attitude toward assistive virtual agents: the role of voice and gender", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 12 No. 4, pp. 4429-4436, doi: 10.1007/s12652-019-01423-x.

Favaro, A. (2019), "Digital culture: fears, hopes and uncertainty about artificial intelligence in Sci-Fi movies", published on ResearchGate [Preprint], available at: www.researchgate.net/publication/335526445_Digital_Culture_Fears_Hopes_And_Uncertainty_About_Artificial_Intelligence_In_Sci-Fi_Movies (accessed 15 November 2021).

Fetzer, J.H. (1990), "What is artificial intelligence?", *Studies in Cognitive Systems*, pp. 3-27, doi: 10.1007/978-94-009-1900-6_1.

Galitsky, B. (2020), "Adjusting chatbot conversation to user personality and mood", *in Human-Computer Interaction Series*, pp. 93-127, doi: 10.1007/978-3-030-61641-0_3.

Hendler, J.A. (2008), "Avoiding another AI winter", *IEEE Intelligent Systems*, Vol. 23 No. 2, pp. 2-4, doi: 10.1109/mis.2008.20, available at: www.emerald.com/insight/content/doi/10.1108/JSM-01-2021-0006/full/html

Jentzsch, S. (2023), "ChatGPT is fun, but it is not funny! humor is still challenging large language models", Cornell University archives, available at: https://arxiv.org/abs/2306.04563

Jertfelt, G.I.H., Blanchin, A. and Li, S. (2016), "Cultural perspective in open-ended interviews – the importance of being adaptable", *Culture & Psychology*, Vol. 22 No. 4, pp. 483-501, doi: 10.1177/1354067X1665080.

Klaus, P. and Zaichkowsky, J. (2020), "AI voice bots: a services marketing research agenda", *Journal of Services Marketing*, Vol. 34 No. 3, pp. 389-398, doi: 10.1108/jsm-01-2019-0043.

Letheren, K., Jetten, J., Roberts, J. and Donovan, J. (2021), "Robots should be seen and not heard … sometimes: anthropomorphism and AI service robot interactions", *Psychology & Marketing*, Vol. 38 No. 12, pp. 2393-2406, doi: 10.1002/mar.21575.

Liu, K., *et al.* (2016), "Two techniques for assessing virtual agent personality", *IEEE Transactions on Affective Computing*, Vol. 7 No. 1, pp. 94-105, doi: 10.1109/taffc.2015.2435780.

Maclure, J. (2019), "The new AI spring: a deflationary view", *AI & Society*, Vol. 35 No. 3, pp. 747-750, doi: 10.1007/s00146-019-00912-z.

McStay, A. (2019), "Emotional AI and EdTech: serving the public good?", *Learning, Media and Technology*, Vol. 45 No. 3, pp. 270-283, doi: 10.1080/17439884.2020.1686016.

Mori, M. (2012), "The Uncanny valley: the original essay by Masahiro Mori", *IEEE Spectrum*, 9 February, available at: https://spectrum.ieee.org/the-uncanny-valley

Peres, R., *et al.* (2023), "On ChatGPT and beyond: how generative artificial intelligence may affect research, teaching, and practice", *International Journal of Research in Marketing*, Vol. 40 No. 2, pp. 269-275, doi: 10.1016/j.ijresmar.2023.03.001.

Picard, R.W. (2000), *Affective Computing*, book published by the MIT press, doi: 10.7551/mitpress/1140.001.0001.

Potdevin, D., Sabouret, N. and Clavel, C. (2020), "Intimacy perception: does the artificial or human nature of the interlocutor matter?", *International Journal of Human-Computer Studies*, Vol. 142, doi: 10.1016/j.ijhcs.2020.102464.

Purington, A., *et al.* (2017), "Alexa is my new BFF': social roles, user satisfaction and personification of the amazon echo", *CHI EA '17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems [Preprint]*, doi: 10.1145/3027063.3053246.

Ramadan, Z., Farah, M.F. and Essrawi, L.E. (2020), "From amazon.com to Amazon.love: how Alexa is redefining companionship and interdependence for people with special needs", *Psychology & Marketing*, Vol. 38 No. 4, pp. 596-609, doi: 10.1002/mar.21441.

Saffari, E., *et al.* (2021), "Does cinema form the future of robotics?': a survey on fictional robots in sci-fi movies", *SN Applied Sciences*, Vol. 3 No. 6, doi: 10.1007/s42452-021-04653-x.

Samuel Craig, C. and Douglas, S.P. (2001), "Conducting international marketing research in the twenty-first century", *International Marketing Review*, Vol. 18 No. 1, pp. 80-90, doi: 10.1108/02651330110398413.

Schweitzer, F., *et al.* (2019), "Servant, friend or master? The relationships users build with voice-controlled smart devices", *Journal of Marketing Management*, Vol. 35 Nos 7/8, pp. 693-715, doi: 10.1080/0267257X.2019.1596970.

Wang, P. (2008), "What do you mean by 'AI'?", *2008 Conference on Artificial General Intelligence*, pp. 362-373, available at: http://cis-linux1.temple.edu/~pwang/Publication/AI_Definitions.pdf

Appendix

Table A1 Main characteristics for participants in population 1

Population 1 experts main characteristics				
Nb.[1]	**Pseudo**	**Sex**	**Professional status**	**Place of residence**
1	Professor N	Man	Doctor in artificial intelligence (AI) and professor of computer science	Paris, *France*
2	Director Y	Man	Head of France, senior advisor in a major tech company with conversational AI	Paris, *France*
3	Founder and CEO H	Man	Co-founder and CEO of a counselling and teaching institute in conversational AI	Copenhagen, *Denmark*
4	Leading conversational AI creator, Dr[2] L	Man	Co-creator of a major conversational AI, CTO and senior vice president of innovation in a major tech company	California, *USA*
5	Director M	Man	Director of innovation and business development in a major tech company and advisor in digital transformation, big data analytics and data science	Paris, *France*
6	Professor T	Man	Co-appointed member of a research centre in AI and professor	Tokyo, *Japan*
7	Founder and CEO B	Woman	Ambassador for women in AI, CEO and co-founder of a start-up in conversational AI	Sydney, *Australia*
8	Dr J	Man	Faculty member, research chair in ethics for AI and robotics, assistant professor	Ottawa, *Canada*
9	Dr S	Man	Research engineer in a major tech company. His research focuses on natural language processing (NLP), deep learning, language generation and textual inference	New-York, *USA*
10	Dr A	Man	Research associate in logic of computation and argumentation in a department of computer science	London, *UK*

Notes: [1]Number given to each interviewee by chronological order; [2]doctor
Source: Authors' own work

Table A2 Main characteristics for participants in population 2

Population 2 quasi-experts main characteristics				
Nb.	**Pseudo**	**Sex**	**Professional status**	**Place of residence**
1	CTO[3] P	Man	CTO (chief technical officer) in a tech start-up with conversational AI	Johannesburg, *South Africa*
2	Senior journalist B	Man	Journalist specialised in new technologies for a major journal	New-York, *USA*
3	Engineer N	Woman	Research engineer and IT developer	Lyon, *France*
4	Director V	Man	Head of growth (digital marketing) in a tech company	Paris, *France*
5	Engineer A	Man	Financial engineer in a major bank group and lecturer in digital technology	Paris, *France*
6	Teacher C	Woman	Teacher in economics and law in an IT department	Lyon, *France*
7	Doctor F	Man	Director of a hospital centre with prior experience of deploying AIs in hospitals	Pontoise, *France*
8	Consultant L	Man	Consultant in information systems and AI	Namur, *Belgium*
9	Manager A	Man	IoT application manager	Lyon, *France*
10	Professor M	Woman	Professor of computer science, research topics in AI	Paris, *France*
11	PhD candidate A	Woman	Doctoral student on gender bias in conversational AI and co-host of a podcast entitled "let's chat ethics"	London, *UK*

Note: [3]Chief technical officer
Source: Authors' own work

Table A3 Main characteristics for participants in population 3

Population 3 non-experts main characteristics						
Nb.	**Pseudo**	**Sex**	**Age**	**Ethnicity**	**Professional status**	**Place of residence**
1	Celine	Woman	21	French *Caucasian*	Finance student	Bouville, *France*
2	Kevin	Man	23	French, Vietnamese *East Asian*	Engineer and consultant	Saint-Maur-des-Fossés, *France*
3	Mina	Woman	22	Madagascar *Sub-Saharan African*	Student in marketing	Antananarivo, Madagascar
4	Ruelle	Woman	23	Chinese *East Asian*	Student employee	Paris, *France*
5	Marc	Man	40	French, Algerian *North-African*	Private driver	Alger, *Algeria*
6	Noemie	Woman	21	French *Caucasian*	Student in management	Darmstadt, *Germany*
7	Lola	Woman	23	French *Caucasian*	Literature student	Paris, *France*
8	Isabelle	Woman	23	Italian *Caucasian*	Marketing student	Issy-les-Moulineaux, *France*
9	Maelle	Woman	22	French *Caucasian*	Marketing student	Neuilly-sur-Seine, *France*
10	Anastasia	Woman	24	French *Caucasian*	Marketing student	Villepinte, *France*
11	Anna	Woman	22	Ethiopian *Sub-Saharan African*	Student working in information systems	Paris, *France*
12	Joelle	Man	26	French, Lebanese Guyanese *Caucasian (West Asian)*	Line driver	Cayenne, *French Guyana*
13	Louis	Man	23	French *North-African*	Student in biology	Les Clayes-sous-Bois, *France*
14	Mike	Man	27	French *Caucasian*	Video editor	Antony, *France*
15	Léna	Woman	23	French *East Asian (China)*	Student working in information systems	Villepinte, *France*
16	Arnaud	Man	27	French *Caucasian*	Entrepreneur in personal development	Blanc-Mesnil, *France*
17	Maria	Woman	22	French *Caucasian*	Marketing student in apprenticeship	Bussy-Saint-Georges, *France*
18	Bastien	Man	18	French and Algerian *North-African*	Business school student	Tokyo, *Japan*
19	Estienne	Man	38	French *North-African*	IT project manager	Paris, *France*
20	Ugo	Man	30	French, *Caucasian (Eastern European)*	Architect	Creteil, *France*
21	Georges	Man	22	French *Caucasian*	Singer and actor	Esbly, *France*
22	Gaspard	Non-binary	25	French *Caucasian*	Job seeker	Nanteuil-les-Meaux, *France*
23	Sarah	Woman	25	French, Algerian *North-African*	Development and international admissions manager	Meaux, *France*
24	Ronald	Man	27	French *Caucasian*	Art gallery student	Paris, *France*
25	Estelle	Woman	33	Franco-Portuguese *Caucasian*	Childcare worker	Champigny-sur-Marne, *France*

Source: Authors' own work

Summary of the third table

Breakdown by gender: 52% (13 people) interviewed were female, while 44% (11 people) were male and one person identified as non-binary.

Age distribution: The majority of interviewees were aged between 20 and 25, i.e. 64% (16 people), 16% of interviewees (4 people) were aged between 26 and 31, while 12% were aged between 32 and 40 (3 people). One interviewee was under 20.

Breakdown by ethnicity: Most of the interviewees were French, but several were of foreign origin or of dual citizenship. In terms of ethnicity, 60% of interviewees were Caucasian (15 people), 12% East Asian (3 people), 20% North African (5 people) and 8% Sub-Saharan African (2 people).

Breakdown by professional status: The majority of women surveyed were students with different statuses: two women were full-time employees, three were students in apprenticeships, one was in a gap year and six were full-time students (i.e. around 53% of the women surveyed were full-time students). By contrast, the majority of men interviewed were either employed or job seekers (only two were students).

Breakdown by place of residence: The place of residence is fairly diversified between interviewees of the third population although, most resided in the Ile-de-France region at the time the interviews took place.

Corresponding authors

Daphne Greiner can be contacted at: Daphne.Greiner@univ-paris1.fr and **Jean-François Lemoine** can be contacted at: jean-francois.lemoine@univ-paris1.fr