‘Affection of a chatbot to create hedonic experiences’

An experiment in a customer service setting

Name: Niek Draaijer

Student number: 1039770

Education: Master Business Administration

Specialization: Innovation & Entrepreneurship

Faculty: Nijmegen School of Management

Institute: Radboud University

Supervisor: K. Sidaoui

Second examiner: dr. P.M.M. Vaessen

Date: 23-07-2021

Table of content

Abstract

Chatbots in the service industry are associated with practically instead of thrilling and delightful experiences. However, certain situations like accidents ask service providers to go beyond offering practical experience by expressing empathy. This study examines if high-level affective expressions of a chatbot are perceived as affective empathic by the participants and to which degree this will lead to a hedonic experience and high customer satisfaction score in sequence. The research included Dutch participants above 18 years old. These participants filed a damage claim of a fictive rafting accident via a chatbot that expressed either low-level affection or high-level affection. The participants evaluated their conversations via a survey. The results indicated strong support for the serial mediation model. Meaning that high-level affective chatbots positively affect customer satisfaction, and perceived affective empathy and hedonic experience positively mediates this relationship. Moreover, perceived affective empathy and hedonic experience are not the only predictors of customer satisfaction. Conversational errors negatively impact customer satisfaction. Furthermore, utilitarian experience is a stronger predictor of customer satisfaction than hedonic experience. Both the low-level as the high-level affective chatbot provided a high utilitarian experience which explains why both chatbots had a marginal difference in customer satisfaction scores in favor of the high-level affective chatbot. This research implies that chatbots are perceived as affective entities. However, companies should create chatbots that deliver a profound utilitarian experience at a minimal level before implementing high-level affective expressions.

1. Introduction

The role of chatbots is becoming more prominent in various social settings such as schools, hotels, shops, and care homes (Broadbent, 2017). Chatbots are valuable and cost-effective means to assist customers (Ferrara et al., 2016). Society uses chatbots more frequently, which is reflected in the expected market value growth from $17.17 billion in 2020 to $102.29 billion by 2026 (Mordor Intelligence, 2021). A chatbot is an automated conversational agent that interacts with its users (Klopfenstein et al., 2017). Technology at hand, e.g., Artificial intelligence (AI), machine learning, and Natural language processing (NLP), are driving chatbots to become more intelligent (Ayanouz et al., 2020). These technologies make the transition of a single task domain chatbot towards a more human-like conversation chatbot imminent (Zamora, 2017). However, this raises questions on the preferred interaction of customers and the desired role of empathy. The following contradictory theories best capture these questions: the computers are social actors (CASA) and the Uncanny valley of mind (UVM). The primer theory advocates that humans mindlessly apply the same social rules in a Human-chatbot interaction (HCI) as in a human-human interaction (HHI) (Isbister & Nass, 2000; Nass & Moon, 2000). The latter theory claims that affective expressions are perceived as uncanny since these expressions are distinctive attributes for humans alone (Gray et al., 2007; Stein & Ohler, 2017).

Nonetheless, empathy is an essential element within the service industry context, especially empathy shown by front-desk employees (Wieseke et al., 2012). Empathy is our reaction to the observed experiences of others (Davis, 1983). It contains affective empathy, which reflects emotional responses to the affective state of others, and cognitive empathy, which reflects the process of understanding another's reaction (Davis, 1983; Eisenberg & Miller, 1987; Mehrabian & Epstein, 1972). When a customer contacts an insurance company with a claim related to a topic such as damaged property or theft, it is apparent that a customer appreciates being understood and receiving an appropriate response by the chatbot. Hence, the emerging role of empathic chatbots fuels the conflict between CASA and UVM and their associated customer experience.

Within an HCI, measuring empathy is dissimilar to HHI. Humans cannot detect an internal state of chatbots due to the different physiology and conscious awareness between humans and chatbots and the lack of internal morality of a chatbot. (Balkenius et al., 2016; Glaskin, 2012a; Paiva et al., 2017). Złotowski et al. (2016) measured the perceived empathy of humans towards chatbots with the following single statement: *"I think the robot* *understands my feelings"* rated on a Likert scale. This statement reflects the cognitive element of empathy by reflecting on the accuracy of recognizing another's ongoing innermost experience, state, or motivation empathy (Davis, 1983). Similar scales do not reflect the depts of empathy within HCI. Additionally, this statement assumes that humans believe that chatbots can detect their innermost state, while these chatbots do not possess an internal morality themselves. Affective empathy, a distinct but related construct (Davis, 1983), reflects the temporary identification with another person's emotional state through an other-oriented emotional response. Since language detects a person's mood, a chatbot can generate an appropriate response by considering a person's affective state. Through expressed affection, the thoughts, feelings, and sensations will provide thorough information on the total experience of a customer-firm interaction (Lemon & Verhoef, 2016; Schmitt, 1999; Sidaoui et al., 2020).

In essence, a great customer experience relates to a person's feelings (Shaw & Ivens, 2002). Customers describe these experiences as": "*I felt like she understood what I wanted, he cared about me, they did everything they could to help, they made me feel I was the most important person in the world"* (Shaw & Ivens, 2002). There are 200 million references about customer experience in cyberspace, and only a few companies truly understand its meaning (Jain et al., 2017). One explanation might be that companies do not fully control customer experience due to its subjective nature (Verhoef et al., 2009). Moreover, marketing literature focuses on executive actions and results rather than the underlying antecedents and consequences of customer experience (Verhoef et al., 2009). One way to grasp the meaning of customer experience is by attributing feelings towards the hedonic experience (Voss et al., 2003). The hedonic experience describes the affective experience of the customer (Schmitt, 1999) and regards the moods, emotions, and hedonic values (Sidaoui et al., 2020). Due to the fragmented literature, concentrating on one experience element (hedonic experience) provides valuable information (Kranzbühler et al., 2018; Palmer, 2010; Sidaoui et al., 2020).

1.1 Research objective and question

Interactions between customers and companies can generate valuable emotional relationships (Gentile et al., 2007). Regarding the customer service setting, we see that swiftness and ease of use are the essential characteristics of chatbots(A. Chopra, 2020; Helpshift, 2019), reflecting a more utilitarian experience than a hedonic experience (Voss et al., 2003). However, previous research had shown that perceived affection from online avatars or chatbots increases pleasure and arousal (Wang et al., 2007), reduces social distances (Keeling et al., 2010), and mitigating the effect of social exclusion (de Gennaro et al., 2020). For this reason, it is valuable to discover to which degree the expressed affection of chatbots are being perceived as affective empathy and subsequently leads to hedonic experiences in a customer service setting. Moreover, it is valuable to assess if higher hedonic experiences increase the satisfaction scores of individuals.

Therefore, the objective is to demonstrate that chatbots in the customer service setting 'submitting a damage claim' create meaningful hedonic experiences through perceived affective empathy originated from affective expressions by chatbots. The affective expressions consist of a chatbot that expresses low-level affection versus a chatbot that expresses high-level affection and their effects on perceived affective empathy, hedonic experience, and customer satisfaction accordingly.
The related research question is: *To which extend do high-level affective expressions of a chatbot in a customer service setting lead to an increase in perceived affective empathy, hedonic experiences, and customer satisfaction accordingly compared to low-level affective chatbot expressions?*

Relevance

There are signs of skepticism towards transforming everyday life robots (Bartneck & Reichenbach, 2005), caused by the limited ability to interact naturally with humans (Wiese et al., 2017). Moreover, current theoretical models treat humans as truly rational and cognitive beings who rationalize every decision (Conner & Armitage, 1998; Moon et al., 2017; Nejad et al., 2004). Furthermore, there is still limited empirical research on the impact of robots on social cognitive processing, development, and well-being in the academic field.

Since chatbots will have a more dominant role in our society, a better understanding of social interactions such as social schemes, behavior, and expressed empathy between humans and chatbots is imperative for social-cognitive development and the quality of interactions in general (Abubshait et al., 2021; Hortensius & Cross, 2018; Paiva et al., 2017; Wykowska et al., 2016). Specifically, this research attributes to the academic field by providing evidence if expressed chatbot affection leads to perceived affective perceptions of humans towards chatbots in creating a meaningful hedonic experience. This attribution advocates the earlier mentioned UVM theory or the CASA theory. Moreover, the role of hedonic experience, which relates to customer happiness and quality of life, needs to be explored further (Jain et al., 2017). Additionally, this research contributes to the experiential marketing field, which advocates researching more intangible elements linked to customers' emotions (Gentile et al., 2007).

This study will contribute to the practical field by assessing the experience of humans towards chatbots affection and the degree to which humans perceived this expressed affection. This information is beneficial for developers of humans like robots (Broadbent, 2017). Moreover, this research should break the barrier of the utilitarian mindset within the service industry (Brandtzaeg & Følstad, 2017; Rzepka et al., 2020) by providing evidence that chatbots can create meaningful relationships and hedonic experiences in this context. Lastly, this research finds evidence for the service industry that creating a hedonic experience is mutually beneficial for customers by assessing the customer satisfaction scores of the customer experience.

1.2 Outline of the thesis

In the introduction, the research context, the research objective, and its relevance had been discussed. The systematic literature review on relevant concepts related to HCI perceived affective empathy, hedonic experience, and customer satisfaction will be discussed in the continuation of the thesis proposal. Subsequently, these relationships are summarized and displayed in a conceptual model together with the associated hypotheses. The methodology section entails the following aspects: the research strategy, the construct measurement, data preparation, the systematic literature review, and the research ethics. Lastly, the results and discussion section concludes the thesis.

2. Systematic Literature Review

This systematic literature review structure is: firstly, we will unravel the meaning of empathy and specifically perceived affective empathy. Hereafter, the concept of customer experience is untangled by expanding on the difference between hedonic experience and utilitarian experience. These concepts, as mentioned above, give us the essential foundation in understanding the most vital concepts within this research. Hereafter, the role of expressed affection of chatbots towards humans and the role of affection in the service industry gives more contextual meaning to HCI and the insurance setting this study seeks to replicate. Henceforth, customer satisfaction will be discussed as the appropriate measure for valuating customer experience. Ultimately, these concepts are discussed and brought together in hypothesis and the accompanied conceptual model.

2.1 Perceived Affective Empathy

Empathy or Einfühlung (feel into) was a concept invented in Germany in the late nineteenth century (Rodrigues et al., 2009). Empathy is our reaction to the observed experiences of others (Davis, 1983). There is no sufficient definition for scientific inquiry (Reniers et al., 2011) due to the ongoing scientific discussion on empathy as a concept of recognizing emotions, experience emotion, or both. Empathy involves two sub-processes: sharing others' internal states and explicitly considering those (Davis, 1983; Eisenberg & Miller, 1987; Mehrabian & Epstein, 1972). The former term is defined as affective empathy and refers to an observers' emotional responses to the affective state of others (Davis, 1983). The latter term is defined as cognitive empathy and reflects understanding another's person reaction (Davis, 1983). Both components are an integral part of empathy as distinct but related constructs (Davis, 1983). The following concepts will be defined in the next section to clarify different terminologies in the empathy literature: high levels of empathy, perceived empathy; sympathy; and affective empathy.

A high level of empathy regards the sum of understanding someone's feelings and the ability to differentiate that this feeling relates to others instead of themselves (Tan et al., 2019). Individuals most likely to respond emphatically possess either high empathic abilities or an emotional attachment to a particular issue (Bendapudi et al., 1996). Perceived empathy is the absolute difference between a person's personal view on an aspect and how they think others view this aspect (Cramer & Jowett, 2010). Empathy differs from the concept of sympathy since the latter relates to the general commitment instead of sharing what another is feeling (Costa et al., 2004).

As beforementioned, affective empathy refers to an observer's empathic response to the emotional responses to the affective state of others (Davis, 1983). Affective empathy is the shared affection or the temporary identification with another's person's emotional state. Perceived affective empathy is the perception that someone else partakes in the same feeling as the other persons' experience at that moment (Bachelor, 1988). Sharing the same emotions is not, per definition, necessary. Namely, Feshbach (1975) refers to affective empathy as emotional responsiveness and, therefore, the comprehension of a broader range of emotions rather than the same emotions. Batson and Shaw (1991) suggest that sympathy, compassion, and tenderness concepts include emotional reactions. Batson et al. (1995) described empathy as an other-oriented emotional response congruent with the perceived welfare of another person.

2.2 Hedonic Experience

Customer experience evaluates the expectation and the stimuli coming from the interaction (Gentile et al., 2007). Customer experience originates from a set of interactions between a customer and the company. This experience is strictly personal and implies different levels of involvement: rational, emotional, sensorial, physical, and spiritual involvement (Gentile et al., 2007; Schmitt, 1999). Moreover, its evaluation depends on someone's expectation and the stimuli coming from the interaction with the company and its offering in correspondence of the different moments of contact or touch-points(Gentile et al., 2007). Schmitt (1999) viewed humans as rational and emotional beings who are concerned with achieving pleasurable experiences. In this line of thinking, he formulated five strategic experiential modules (SEMs) to capture the total experience. These SEMs consist out of the following dimensions: '*Sensory Experiences (SENSE); affective experiences (FEEL); creative cognitive experiences (THINK); physical experiences, behaviors, and lifestyles (ACT); and social-identity experiences that result from relating to a reference group or culture (RELATE).'* Schmitt (1999) and Gentile et al. (2007) concluded that both the affective and cognitive parts of the experience occur. Well-used models such as the technology acceptance models (TAM), theory of planned behavior (TPB), and theory of reasoned action (TRA*)* try to explain customer behavior (Moon et al., 2017). However, these models indirectly suggest that people are purely rational in their decision-making; hence they are cognitive beings. Since these models explained the affective experience to a limited extent (Moon et al., 2017), it had been proposed to include more affective variables into these models (Conner & Armitage, 1998; Nejad et al., 2004).

(Past) experiences are forming the hedonic attitudes of customers (Voss et al., 2003). Combined with utilitarian attitudes, hedonic attitudes are an integrative two-dimensional conceptualization of consumer attitudes (Voss et al., 2003). Attitudes are emotionally infected filters needed to rearrange the chaotic environment into one an individual can understand (Shrigley et al., 1988). The elements of attitudes consist of cognition, affection, and conation (Fishbein & Ajzen, 1977). Research on attitudes in diverse disciplines such as sociology, psychology, and economics (Batra & Ahtola, 1990). Customer experience had been explored in many components of product experience, like physiological arousal, affect in general, involvement, product satisfaction resulting in specific cognitions, and global product evaluation (Diefenbach & Hassenzahl, 2011).

The utilitarian experience describes the feelings towards the functionality of an object or product (Batra & Ahtola, 1990; Voss et al., 2003). Utilitarianism regards practicality, efficiency, helpfulness, and functionality (Voss et al., 2003). The hedonic experience refers to the sensation felt towards an object or a product (Voss et al., 2003). Therefore, the utilitarian customer experience regards the practical, task-specific, and economic aspects of a product or service (Overby & Lee, 2006). In a chatbot setting, motivations of chatbot usage of customers relate to utilitarian aspects such as productivity, efficiency, and convenience (Brandtzaeg & Følstad, 2017; Rzepka et al., 2020). In the mobile internet services (MIS), which contains customer banking and text messaging (Khedhaouria & Beldi, 2014), perceived usefulness values were higher than perceived enjoyment values. Therefore, MIS was related to a more utilitarian experience than hedonic experience (Khedhaouria & Beldi, 2014). This research hints that context is vital to determine the weight of utilitarian versus hedonic experience.

Hirschman & Holbrook (1982) defined hedonic consumption as: '*facets of consumer behavior that relate to the multisensory, fantasy and emotive aspects of product usage experience.*' Multisensory implies the receipt of more than one mode of experience (taste, sound, scent). Besides the receptive part, hedonic contains internal multisensory imagery (recalling something or fantasizing about something) and emotional arousal (fear, rage, joy) (Batra & Ahtola, 1990). Hedonic consumption is theoretically tied to various behavioral sciences like sociology, aesthetics, and psychology before being embedded in the marketing theory (Hirschman & Holbrook, 1982). Due to its imaginative nature, hedonic consumption is part of a self-constructed reality, which means that questions beyond the objective context are essential to understand each reality (Hirschman & Holbrook, 1982). One way to achieve this is to ask the consumer if they ever pretend that something else is happening during the experience, like consuming a product (Swanson, 1978). Hedonic experiences capture the uniqueness of a product or service or the emotional connection evoked in the consumer (Overby & Lee, 2006). Chatbots can evoke hedonic elements such as entertainment. The motivations of using chatbots as a communication channel relate to hedonic elements such as passing the time, novelty, and social aspects (Brandtzaeg & Følstad, 2017; Rzepka et al., 2020).

2.3 Expressed Affection in a Human Chatbot Interaction

A chatbot needs a high degree of autonomy to express and simultaneously evoke empathy. However, virtual agents are currently tied to specific scenarios and systems, challenging generalizability (Paiva et al., 2017). Empathic processes need to be simulated computationally for social agents to exhibit empathic behavior (Glaskin, 2012b). These processes include empathy mechanisms, empathy modulation, and empathic responses. Empathy mechanisms include recognizing the affective state, understanding a person's actions, and adapting the affective state towards the human affective state. The perception of context and reasoning about the aspects at hand embodies the cognitive state of the empathy mechanisms (Paiva et al., 2017). Empathy modulation regards the modulation of relationships, personality, and moods (Paiva et al., 2017). Affective empathy regards the expression of the internal state in an understandable way for others and performing actions that change the state of the other. (Paiva et al., 2017). It has been contended that only intrinsically moral chatbots can establish the trust of humans by both the users and the public at large (Balkenius et al., 2016).

Due to different physiology and conscious awareness, the emotional experiences between humans and robots are distinct (Glaskin, 2012). Therefore, the empathy of social agents is measured from a human perspective through perceived empathy (Paiva et al., 2017). Schmetkamp (2020) perceived empathy in Human-Robot-Interaction (HRI) as a process and not so much as an outcome. Moreover, Schmetkamp (2020) found that empathy in HRI can minimally be established on an imaginative perspective-taking level by humans. This evidence is reflected in various studies (Cross et al., 2019a; de Gennaro et al., 2020; Konijn & Hoorn, 2020; Liu & Sundar, 2018; Rosenthal-Von Der Pütten et al., 2014). To exemplify, the emotional responses of an empathic chatbot have a mitigating effect on the feeling of being socially excluded in the case of cyberbullying (de Gennaro et al., 2020). Demonstrating high levels of emotional communication does not always result in high levels of perceived empathy. For example, humans perceived computer-controlled agents as eerier when these entities were perceived as highly empathic (Stein & Ohler, 2017). People are cautious when chatbots replicate or resemble the human brain with empathic responses (Stein & Ohler, 2017). Furthermore, negatively framed messages from chatbots which are perceived as criticism, enlargers the level of anxiety of humans towards robots (Złotowski et al., 2016). Additionally, perceived empathy does not permanently alleviate anger or negative responses when a robotic entity makes a mistake (Chen et al., 2021). Moreover, the appropriateness of responses of conversational agents (Siri, Cortona, e.g.) such as empathic, informative, directive, and confirmatory responses dropped when questions were rephrased (Kocaballi et al., 2020). This drop symbolizes the lack of autonomy, capabilities, and intrinsic morality needed to create a more human-like conversation (Balkenius et al., 2016).

Another important notion is the visual impact of chatbots on humans. The facial design affects perceived sociability (Rossi et al., 2020). In the customer service industry, it is not inappropriate for robots to show facial emotions in settings where this is needed (Konijn & Hoorn, 2020). Artificial pictures did not decrease perceived empathy towards the digital persona (Salminen et al., 2020). Humanlike robots decrease perceived empathy levels and trustworthiness more than machinelike robots (Złotowski et al., 2016). Thus, facial expressions do not necessarily need to be human. Similar findings are found in the gaming industry for avatars, non-player characters, and in HCI in general (Ho & Ng, 2020; Prendinger et al., 2006; Sierra Rativa et al., 2020; Swiderska & Küster, 2018; Wilde & Evans, 2019).

The duration of interaction influences empathy and experience in an HRI. Rosenthal-Von Der Pütten et al. (2014); Konijn & Hoorn (2020) demonstrated that humans are sensitive to exposed affection or violence in an HRI. However, in a similar setting, short-term interaction (within a week) leads to less perceived empathy in an HRI (Cross et al., 2019a). A call for long-term HRI is needed to explore further the experience in an HRI setting (de Gennaro et al., 2020; Paiva et al., 2017).

Lastly, in the theory of HCI, there are two contradictory theories: CASA and UVM. The primer theory advocates that humans mindlessly apply the same social rules in an HCI as in an HHI (Isbister & Nass, 2000; Nass & Moon, 2000). Prior research has shown that, in general, computer-based supportive agents have been received positively by humans (Brave et al., 2005; Klein et al., 2002; Partala & Surakka, 2004). Hence, according to CASA, affective expressions are embraced and cause no disturbance within a conversation. The latter theory suggests that expressions of feelings and emotions violate people's norms since these expressions are conversation are a distinctive attribute for humans alone. The violation of these norms elicits perceptions of uncanniness (Gray et al., 2007; Stein & Ohler, 2017). This term captures the perceived threat to human distinctiveness.

2.4 Expressed Affection in the Service Industry

Empathy within the service industry is an essential aspect of human interaction and is a critical competence for customer service agents (Clark et al., 2013). Service providers who adapt their behavior to the consumer are perceived as more empathic by these consumers (Collier et al., 2018). In general, the role of empathy in the service sector is an undervalued component compared to tangibles, reliability, assurance, and responsibility (Costa et al., 2004). From an experiential point of view, empathy is a thin line. Too much empathy is unwanted since too much care, and concern lead to 'over-servicing, which negatively affects customer experience (Tan et al., 2019). However, ignoring the affective state or failing to recognize a customer negatively impacts the customer experience (Turley & O'Donohoe, 2017). Another aspect making it harder to balance empathy is the demand for variations of empathy types (Clark et al., 2013). These variations require the service provider to listen, assess and provide appropriately. In call centers, agents need to build rapport but simultaneously maintain an emotional distance to a certain degree. Furthermore, the positioning of service providers by customers plays a role in the desired empathy by these service providers. Customers' empathy is elicited when service providers showed less competent service but had high morale (Kirmani et al., 2017).

In the context of online shopping, feelings of empathy towards an agent are induced by swiftness, clarity, information, ease of access, and personalization (Suryandari & Paswan, 2014).

2.5 Customer satisfaction

Customer satisfaction is: "*The number of customers, or percentage of total customers, whose reported experience with a firm, its products, or its services (ratings) exceeds specified satisfaction goals*" (Bendle et al., 2016). Experiences that fall below the individuals' expectations will lead to a less satisfied customer, and experiences above someone's expectation will lead to a more satisfied customer to a more delightful experience (Oliver et al., 1997). Since one could have a pleasurable experience but still be dissatisfied with the encounter (Yi, 1989), it is valuable to include this variable in this research setting. Customer satisfaction research defines two movements, namely transaction-specific satisfaction and cumulative satisfaction (Johnson et al., 1995). The latter regards the satisfaction level over a more extended period. The primer regards the customer experience of a singular service encounter (Yi, 1989) and is for this reason more fitting in this research context. Customer satisfaction is not an unusual concept in the evaluation of chatbot encounters. Recent customer service studies on chatbot experiences had used customer satisfaction to score the individual assessment on the overall conversation quality (K. Chopra, 2019; Chung et al., 2020; McLean & Osei-Frimpong, 2019).

2.6 Hypotheses and conceptual model

Affective Expressions → Perceived Affective Empathy

A chatbot with an emotional communication style detects and understands the user's emotions and responds to them on an appropriate emotional level (Liu & Sundar, 2018). These empathic expressions of a chatbot are preferable to non-empathic responses (Liu & Sundar, 2018). However, since chatbots are text-based, it is not easy to include (facial) expressions or gestures to match someone's emotional state (Spring et al., 2019). Linguistic elements of a chatbot can be manipulated to display different communication styles (Liebrecht et al., 2021), which might influence how humans identify themselves with chatbots (Xu & Lombard, 2017). The manipulation of language style, for example, results in more friendly perceived virtual customer service agents (Verhagen et al., 2014).

In general, the social reactions of humans increase when computers provide more social cues (Nass & Moon, 2000). For example, human-like cues of a virtual agent positively impact the emotional connection (Araujo, 2018). This emotional connection reflects the feeling of being understood and cared for and indicates a connection with at least perceived affective empathy. Empathic expressions are reflected in the basic emotions of Ekman (1992), which are: happiness, sadness, fear, anger, disgust, and surprise. These basic emotions annotate deep learning algorithms to generate appropriate emotional responses of chatbots towards humans (H. Zhou et al., 2017), such as high-level affective expression (Liu & Sundar, 2018).

The affective expression relates to sentences in which a chatbot expressed that it understood how and why the person feels a certain way (Liu & Sundar, 2018). The expectation is that high-level affective expressions of a chatbot lead to a higher degree of perceived affective empathy of humans towards a chatbot since perceived affective empathy reflects the perception of someone's emotional response to the affective state of this person (Davis, 1983). Therefore, the CASA theory's underlying assumption, indicating that humans apply the same social rules for chatbots. Low-level affective expressions indicate that the chatbot will express itself without perspective-taking or expressing compassion to a great extent towards a person (Liu & Sundar, 2018). Thus, the expectation is that a chatbot with high-level affective expressions leads to higher perceived affective empathy scores compared to a chatbot with low-level affective expressions.

H1a: A chatbot's high-level affective expressions have higher perceived affective empathy scores than low-level affective expressions.

Perceived Affective Empathy → Hedonic Experiences

Employees who expressed affection such as friendliness and caring are more likely to create customer delight (Barnes et al., 2011). Hedonic attitudes are derived from sensational experiences (Voss et al., 2003), whereas firms generate customer delight through hedonic attributes such as experience and enjoyment-related benefits (Chitturi et al., 2008). Therefore, exceeding a customer's expectations through expressed affective empathy results in a more pleasant experience leading to customer delight (Nguyen et al., 2020). Experiences of empathy by customers that contain emotional components are more intense and meaningful (Kerem et al., 2001). The term meaningful is reflected in utilitarian items such as helpfulness and usefulness. However, the term intensity, or' passion' in experiences, is expected to be reflected in hedonic items such as thrilling, exciting, and delightful. In extension, the social signals of chatbots impact the hedonic experience of customers by generating emotional and affective reactions of humans (Etemad-Sajadi, 2014; Holzwarth et al., 2006). Since appealing to a customer's inner feelings and emotions create affective experiences (Schmitt, 1999), the expectation is that the hedonic attitudes towards chatbots will be positive affected by perceived affective empathy.

H1b: An increase in perceived affective empathy scores towards the chatbot directly increases hedonic experience scores of the chatbot interaction.

Hedonic experience → Customer satisfaction

Increasing customers' empathy leads to higher customer satisfaction evaluations in the service provider settings (C. Davis et al., 2017). However, the same study does suggest that empathy most likely increases cooperation of customers in which satisfaction is one way of indicating the cooperation towards a service provider, making it hard to prove a direct link between empathy and customer satisfaction. The expectation is that a more hedonic customer experience directly increases the customer satisfaction scores. Past research indicates that hedonic values and utilitarian values from shopping experiences directly increase customer satisfaction evaluations. In which hedonic values had a more substantial impact than utilitarian values (Babin et al., 1994; Eroglu et al., 2005).

Nonetheless is the context of the experience vital. In a fast-food restaurant setting, utilitarian aspects of the overall experiences impact customer satisfaction more strongly than the hedonic counterpart (Ryu et al., 2010). Elements such as convenience and quick serving should be more critical assessment criteria in this setting (Ryu et al., 2010). However, chatbots are automatically programmed to be convenient and offer immediate service (Brandtzaeg & Følstad, 2017; Rzepka et al., 2020). Hence, these elements are minimum expectations or dissatisfiers (Johnston, 1995). Hedonic experiences such as fun and sensation are unlike convenience and quick service, not expected in an insurance service setting. So when hedonic elements are positively received, the customer satisfaction scores are expected to be impacted positively.

H1c: An increase in hedonic experience scores of the chatbot interaction lead to an increase in customer satisfaction scores.

Comparing the (in)direct effect of affective expressions on customer satisfaction

Past research had shown an indirect effect of empathy on customer satisfaction. For instance, empathy increased service quality, positively affecting customer satisfaction scores (Parasuraman et al., 1991). Feelings of satisfaction in an interpersonal relationship are associated with empathy. This effect for customers had not been researched extensively in a service provider setting (Davis et al., 2017). Nonetheless, in various settings, the positive effect of empathy on satisfaction had been concluded in studies on doctor-patient relationships (Kim et al., 2004); coach-athlete relationships (Jowett et al., 2012); and management- stakeholders relationships (Strong et al., 2001). Expressed empathy leads to a tendency to feel. These expressions are experienced more intensively in a service situation, resulting in greater customer satisfaction (C. Davis et al., 2017). Thus, the relation of affection towards satisfaction is indirect. Therefore, the expectation is that a high level of affective chatbot expressions generates higher scores on perceived affective empathy, which positively influences the hedonic experience and leads to positive customer satisfaction scores compared to a low level of affective chatbot expression. The expectation is that the indirect effect of Ind3 (see figure 1) has a more substantial positive impact on customer satisfaction than the alternative (in)direct effects. The following hypotheses therefore are:

H2a: The indirect effect 3 has a more substantial impact on customer satisfaction than indirect effect 1
H2b: The indirect effect 3 has a more substantial impact on customer satisfaction than indirect effect 2
H2c: Indirect effect is more substantial than the direct effect

Control variables

Past Experience

In an HCI setting, the empathic perception of humans towards chatbots will most likely differ from an HHI setting. Humans have more expertise in observing and interacting with humans than robots, which partly explains why people respond and perceive chatbots differently (Cross et al., 2019b). Moreover, since chatbots do not possess a similar internal state as humans, only a certain degree of alignment of representations is found regarding the perceived empathy of humans towards chatbots. However, chatbots evoke imaginative perspective-taking levels by humans at a minimum (Schmetkamp, 2020). Therefore, in the situation where people are more familiar due to previous interaction with a certain chatbot, it is expected that these persons have better accuracy of the representations of a chatbot compared to people who have had no previous experience or interaction with a chatbot. Previous research indicated that people unfamiliar with the chatbot medium attributed lower scores of friendliness or socialness to the chatbot (Ischen et al., 2020). The expectation is that the more familiar someone is with the chatbot forum due to past experiences, the greater the chance that a human is open to chatbots' expressed affection.

Susceptibility to empathy

Differences in empathic abilities or emotional attachment to a particular issue between individuals lead to a different interpretation of expressed affection (Bendapudi et al., 1996; Tan et al., 2019). One person might be more susceptible to expressed affection than another. The expectation is that respondents who are more susceptible to empathy will attribute more meaning to high-level affective expressions of a chatbot.

Age

As individuals are maturing, the ability to manage emotions increases which subsequently decreases the empathic experience of this person (Batson et al., 1987; Eisenberg & Miller, 1987). Hence, the indirect effect of affective expressions on customer satisfaction is more substantial for a younger person than more senior persons. Moreover, in earlier research of empathy in an HCI setting, age was controlled as a covariate (Liu & Sundar, 2018). The expectation is that the indirect effect of high-level affective expressions on customer satisfaction decreases of participants when age increases. The expectation is that this effect will occur around the age of 40 years.

Gender

In general, persons who identify themselves as females are more emotionally responsive than males and find it easier to empathize with a target (Eisenberg & Lennon, 1983). Therefore, many empathic studies only include female participants (Batson et al., 1997; Batson, Klein, et al., 1995; Batson, Turk, et al., 1995). More recent research conducted in a negative service setting displayed that the inducement of empathy by employees generated only positive significant customer satisfaction results for female customers (C. Davis et al., 2017). Altogether, it is expected that the indirect effect of high-level affective expressions on customer satisfaction will be more substantial for persons who identify themselves as women compared to men.

Attitude

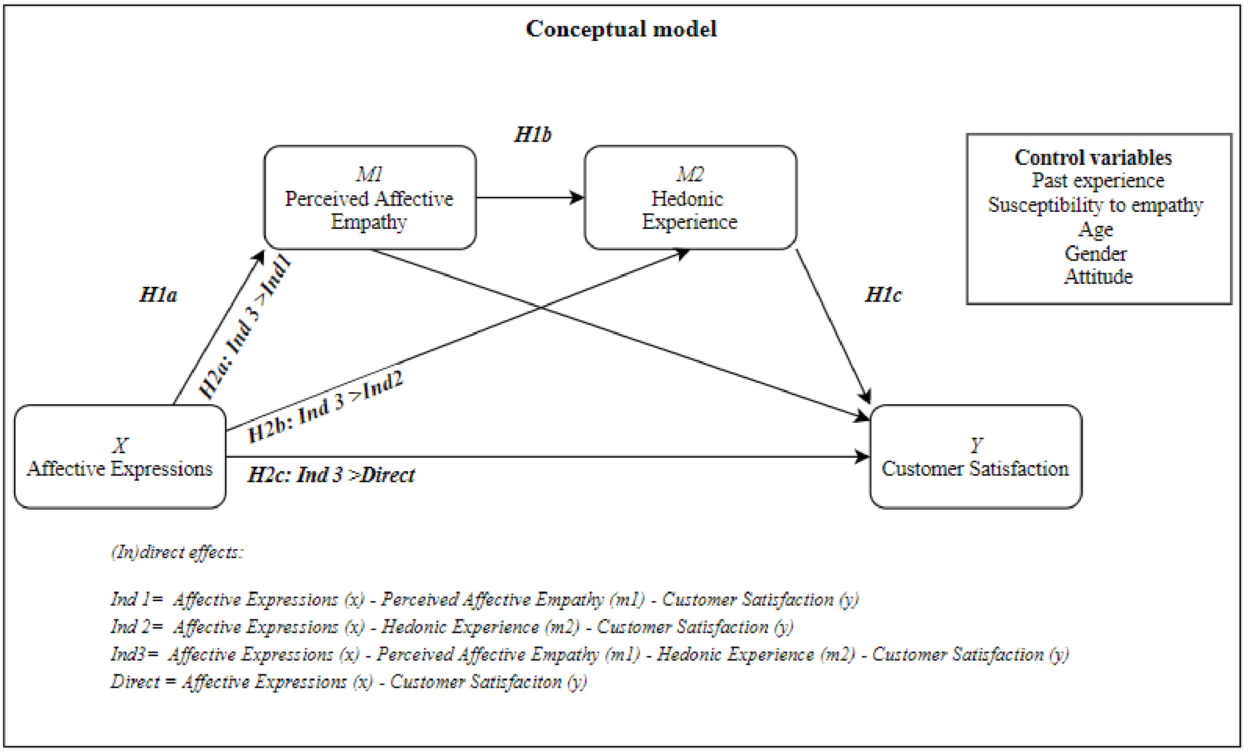
A common way to explain and predict the attitudes of new technologies by users is found in TAM (technology acceptance model) (F. D. Davis, 1989). This model explains that attitudes towards technology are based on the perceived usefulness and perceived ease of use in relevant HCI fields (King & He, 2006). However, chatbots are relatively new to customers. As a consequence, the attitudes of humans towards chatbots had not been researched to a large extent (Rietz et al., 2019), but new research is on the rise. To exemplify, recently, attitudes and customer acceptance of chatbots in the online shopping experience (Araújo & Casais, 2020; Rese et al., 2020). The expectation is that high-level affective expressions' indirect effect on customer satisfaction is more substantial for respondents with a positive attitude towards chatbots.

The conceptual model

Altogether the proposed hypotheses and relations regarding the (in)direct effect of affective expressions on customer satisfaction are illustrated in the conceptual model below (Figure 1).

Figure 1

Conceptual model of empathy on customer experience



3. Methodology

3.1 Research strategy

The conducted research is quantitative and takes shape in the form of a survey. One distinct characteristic of a survey is the inclusion of many comparable objects collected systematically (Vennix, 2019). A large sample is needed in this research setting to statistically find support for the proposed conceptual model and test if the conceptual model is an accurate representation of reality (Vennix, 2019).

Scenario

The participants were introduced to a fictive scenario where they sustained a broken leg during a rafting trip in Switzerland. Fortunately, the participant bought an 'extreme sports insurance,' which disembarks the cost of any injuries during the vacation in Switzerland. For which the participant had to fill in a damage claim via the constructed insurance chatbot.

The chatbot profiles

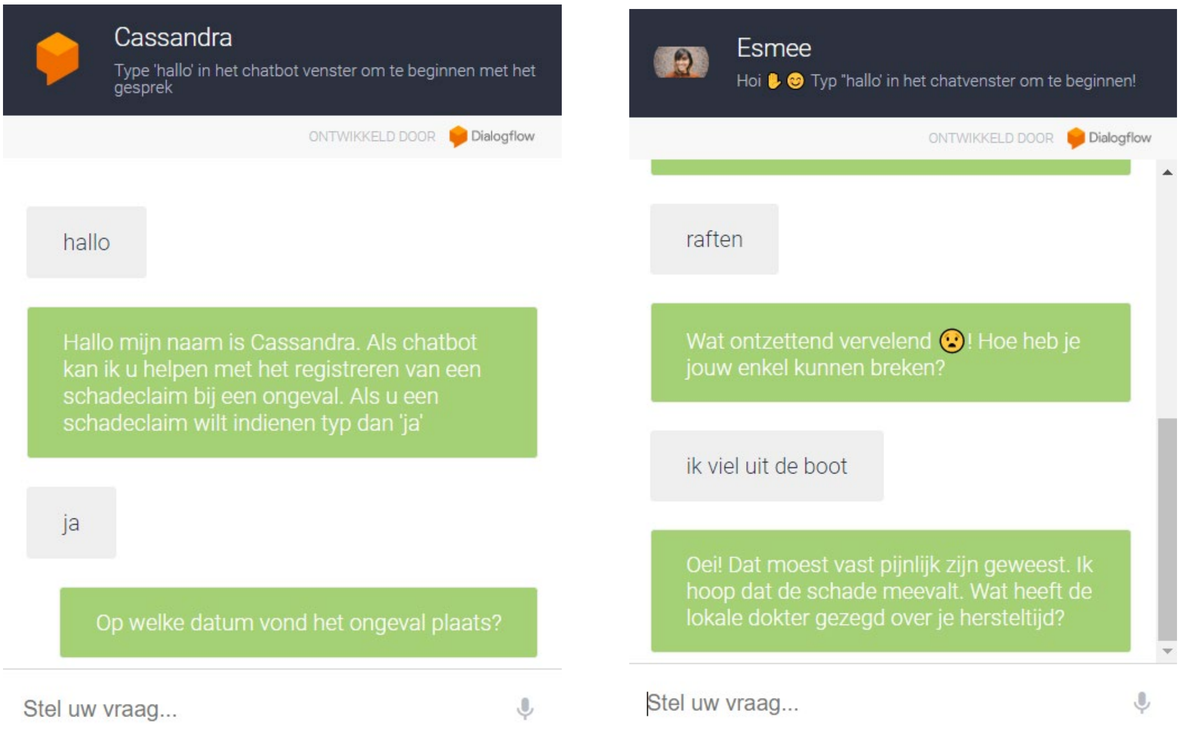
The constructed chatbots are Cassandra and Esmee. Cassandra represents the chatbot with low-level affective expressions and is an efficient chatbot with more neutral emotional responses. Esmee represents the high-level affective expression chatbot and responds more emotionally through affection and empathic expressions via informal language (e.g., emoticons, friendliness, kindness).

Procedure

The procedure relates to an experiment in which a health advice chatbot expressed empathy (Liu & Sundar, 2018). Firstly the participants are introduced to the scenario in Qualtrics as described in the section above. After reading this scenario, the participants were randomly assigned to either the high-level affective chatbot or the low-level affective chatbot via a URL. The participants were to fill in a damage claim with their assigned chatbot. After the task, they were asked to return to the Qualtrics survey to evaluate the conversation. The order of these questions was: Perceived affective empathy; Customer Experience; Customer satisfaction; Additional control variables. Overall the entire survey, including the chatbot interaction, took around 6-10 minutes.

Figure 2

The chatbot profiles



Sample

The population of this research is Dutch citizens ≥ 18 years. Dutch citizens are obligated to acquire basic health insurance when they turn 18 and therefore had (to some degree) encounters with insurance agencies or at least a basic understanding of insurances. To reach this population, the non-probability sampling techniques, convenience sampling, will be used. This sampling technique considers the population at hand but is limited because only probability sampling assumes representativity. Therefore, external validity is not promoted (Field, 2017). Notwithstanding, convenience sampling is a practical, inexpensive, and easy way to collect data.

The desired level of the sample size to conduct analyses via OLS linear regression is established on 15-20 observations per variable (Field, 2017). When this ratio is lower than 5:1, there is a risk of overfitting the variate to the sample. Overfitting harms the generalizability of the results (Field, 2017). Since the conceptual model consists of 11 variables, the desired respondents vary between 165-220, with an absolute minimum of 55. G *power 3.1.9.7, a software program that computes statistical power analyses, is used to double-check the required sample size (Faul et al., 2009). For a medium effect size (f2 = .15), with an alpha of .05, a power of .80, and ten predictors, the total sample size is estimated on 118 respondents.

The results showed that 97 respondents successfully finished the survey. 52% of the survey represented the male population and 48% the female population. Regarding age, 77% of the sample represents the age group 19-27 years. 22,7% of the sample is > 27 years. 44% of the participants claimed they never had a conversation with a chatbot. 54% did claim to have had a conversation with a chatbot. The mentioned chatbots were related to online shopping (Billy from Bol.com and Zalando), logistics (PostNL), travel (Ryanair and NS), and personal services like banking, insurances, and internet providers (e.g., CeeBee of Centraal Beheer, KPN, and ING). 28% of the participants preferred to be assisted by a chatbot. 23% were neutral, and 48% preferred not to be assisted by a chatbot. Lastly, the degree to which the participants are open to a stranger's compassion if they are not feeling well indicated that 55% are open to strangers' compassion, 25% preferred not to receive compassion from a stranger, and around 18% remained neutral.

Analyses settings and methods

Bootstrapping

Bootstrapping estimates the properties of the sampling distribution from the sample data. In other words, the sample data will be treated as a population (bootstrap samples) from which numerous samples are drawn (Field, 2017). By this means, the sampling distribution does not need to be normally distributed and solves the potential problems associated with smaller sample sizes. The bootstrapping procedure is commonly used in the PROCESS software to calculate confidence intervals. Table 1 displays the software programs and the purpose of these programs for this research.

Table 1

Software tools

Software	**Use**	**Source**
Google Dialogflow ES	This program is used for constructing and displaying the chatbot conversation.	(*Dialogflow*, n.d.)
SPSS IBM v.26	This is a statistical analytics software program that conducts research analyses.	(*IBM SPSS Statistics 26*, n.d.)

*ADANCO v.2.2.1	PLS-SEM software for, e.g., Confirmatory Factor Analyses.	(ADANCO, n.d.)
**PROCESS v3.4.1	PROCESS is a software program to estimate mediation models via bootstrapping.	(Hayes, 2012)
Qualtrics	Survey tool licensed by the Radboud University.	(*Qualtrics Online Survey Tool*, n.d.)

**ADANCO*

ADANCO is a software program that allows conducting structural equation modeling (SEM) and performing confirmatory factor analyses using bootstrap-based tests.

***PROCESS*

PROCESS is a software program that specifies models and estimates mediated effects (Hayes, 2012). PROCESS is accessible via the statistic software program IBM SPSS. PROCESS allows the addition of any number of predictor variables such as covariates in a linear equation model (Hayes, 2012). One of the disadvantages of PROCESS regards the estimation procedures of solely the observed variables. PROCESS is less sophisticated and offers fewer challenges compared to other methods such as SEM. Therefore, using this software program reduces potential, which otherwise would have happened (Demming et al., 2017; Iacobucci et al., 2007).

3.2 Construct measurement

In the situation where the independent variable is categorical, and both the mediating and independent variables are continuous, regression analyses are appropriate (Iacobucci, 2012). Moreover, the purpose of this particular research is to expose a serial indirect effect, which matches the foundation of serial mediation (Demming et al., 2017).

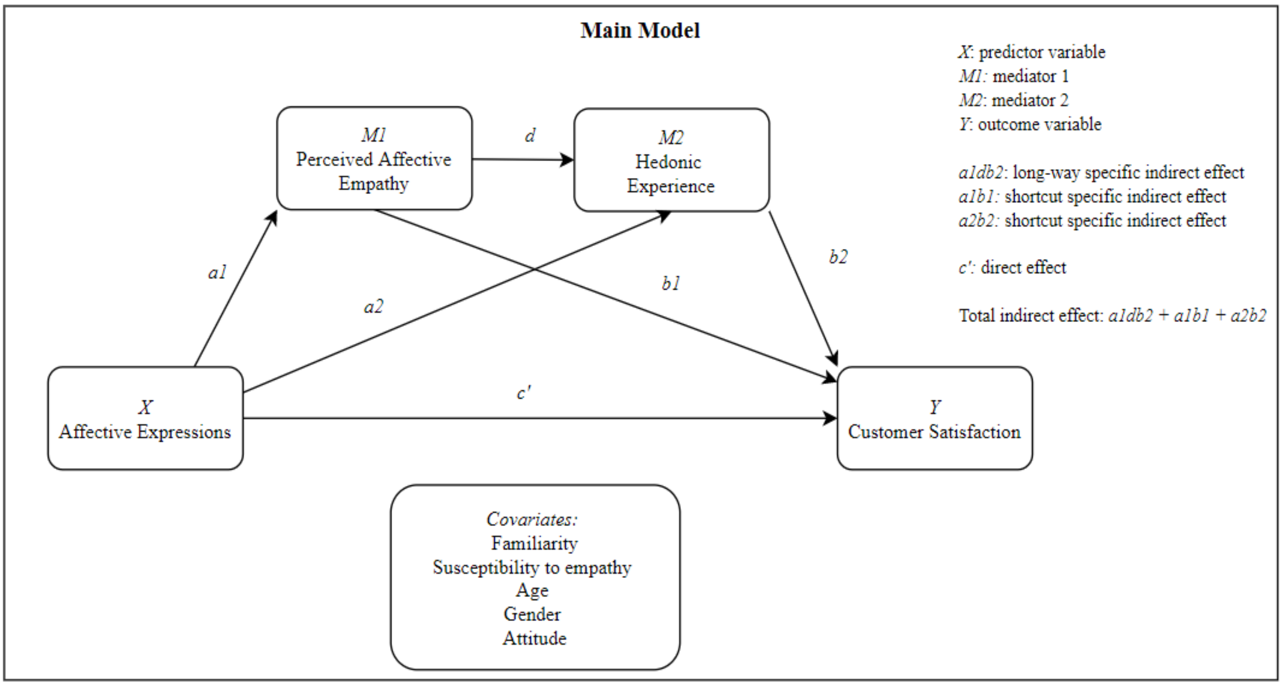
The model: Serial mediation

The proposed serial multiple mediator model assumes a causal chain linking the mediators with a specified direction of causal flow (Montoya & Hayes, 2017). The advantage of using serial mediation compared to a singular mediation model is that the complexities of X's effect on Y are faced more extensively (Montoya & Hayes, 2017). The Demming/Jahn/Boztug typologies for conducting mediation analyses suggested that a serial mediation analysis is the most appropriate way forward (Demming et al., 2017). Namely, the model proposed a mediator. Secondly, none of the model paths are conditional on a moderator. Thirdly, there is more than one mediator proposed. Finally, the expectation is that these proposed mediators are causally related. The multiple indirect effects of the serial mediation model could lead to new insight and open the debate for competitive theories, mainly because the proposed conceptual model is new for this exact research context.

The control variables portrayed in figure 3 will be included as covariates within the serial mediation model and therefore account for any shared associations between variables in the causal system caused by other sources (Hayes, 2012).

Figure 3

Statistical model: Serial mediation



Operationalization

Table 2 includes the operationalization of the variables in the proposed serial mediation model. Hereafter are the scales of perceived affective empathy, hedonic experience, and customer satisfaction more thoroughly discussed in the section 'scales'.

Table 2

Operationalization of the variables

Main model	Measurement description	Measurement level	Author(s)
Affective Expressions	- Low level versus high-level affective expressions	Nominal	(Liu & Sundar, 2018)
Perceived Affective Empathy	- RoPE scale: eight items on a 5-point Likert scale	Interval	(Charrier et al., 2019)
Hedonic Experience	- HED/UT scale: eight items on a 7-point semantic differential scale	Interval	(Heijden & Sørensen, 2003)
Customer Satisfaction	- Satisfaction: one item on a 5-point Likert scale - NPS: one item on an 11-point scale	Interval	(Bendle et al., 2016) and (Reichheld, 2003)
Control variables			
Past Experience	- Previous experience: 'yes' or 'no' plus open coding	Nominal	(Park & Stoel, 2005)
Age	- Median split	Interval	-
Gender	- Identifies as male, female, other	Nominal	-
Attitude	- Attitude: 1 item on a 7-point scale	Interval	Derived from (Davis, 1989)
Susceptibility to empathy	- 1 item on 1 item on a 7-point scale	Interval	-
Conversational errors	-1 item on a self-constructed scale that indicates errors in the conversation	Ordinal	

Scales

Perceived affective empathy

The measurements of perceived affective empathy in an HCI setting are not robust and untested in terms of validity and reproducibility (Charrier et al., 2019). Furthermore, previous measurements were biased because of the presumption of cognitive abilities among robots (Nomura, 2019). The RoPE (Robot's Perceived Empathy) scale overcomes this bias (Charrier et al., 2019). The dimensions of the RoPE scale are empathic understanding which represents perceived cognitive empathy, and empathic response, which measures perceived affective empathy. The focus of this study will be on the empathic response, which measures how one communicates their attempts at understanding someone else's frame of reference, including their feelings and motives. The assessment of the scale is via a 5-point Likert scale running from "does not describe me well" to "describes me very well" (Charrier et al., 2019). The RoPE scale originates from a previous HCI pilot study integrating Davis's IRI (Interpersonal Reactivity Index) (Charrier & Galdeano, 2018). The IRI is a self-reported measurement scale of the reaction from one individual to the observed experiences of another (Davis, 1983).

Hedonic experience

The most used scale to measure hedonic experience is the HED/UT scale (Voss et al., 2003). The HED/UT scale entails the hedonic (HED) and utilitarian (UT) dimensions and measures these dimensions on an eight-item 7-point semantic differential scale. The HED/UT scale is reliable, valid, parsimonious, and generalizable (Bagozzi & Burnkrant, 1979; Batra & Ahtola, 1990; Voss et al., 2003). Moreover, the HED/UT scale is present in the Marketing Scales handbook (Heijden & Sørensen, 2003). Lastly, the scale does not violate convergent validity (Bagozzi & Burnkrant, 1979; Batra & Ahtola, 1990; Voss et al., 2003). The original HED/UT scale of Voss et al. (2003) had been adopted in the MIS setting by Heijden & Sørensen (2003). The latter authors justified this adoption by stating that the founders of the original scale suggested that the HED/UT scale is generally applicable to all products and services (Heijden & Sørensen, 2003).

Customer Satisfaction

It is more beneficial to conduct research in which multi-item customer satisfaction measures. Multi-item measurements have a higher explaining power of customer satisfaction, and they are more reliable than single-item measurements (Yi, 1989). There is not yet a consensus on the perfect satisfaction scale due to the overall skewness caused by the high satisfaction ratings of respondents (Yi, 1989). However, it is wise to use parsimonious scales since an increase in categories does not necessarily improve validity or utility (Peterson & Wilson, 1992). The first item entails the satisfaction scale. This scale measures the degree of satisfaction from very satisfied" to "very dissatisfied" (Bendle et al., 2016). The second item is the Net Promotor Score (NPS). This scale measures the extent to which one person recommended services or brands (Morgan & Rego, 2006). The NPS is a proxy for future behavioral intention and measured on an 11-point scale ranging from zero to ten which varies from zero 'not at all likely' to ten 'extremely likely'(Reichheld, 2003).

3.3 Data preparation and cleaning

Checking data matrix

The data had been converted into SPPS and checked on missing scores via descriptive analysis (e.g., frequencies and crosstabs). These actions ensure that no systematic errors in the survey potentially affect the internal validity (Field, 2017). Due to non-response, 91 of 189 respondents were excluded from the data set. The data indicates that this occurs at 18% of the survey. This percentage indicated the start of chatbot conversation. Therefore almost half of the participants did not want to proceed with the survey because of the chatbot interaction. Moreover, the items; Aff_Emp_1; Aff_Emp_6; Cog_Emp_3; and Cog_Emp_5 were reversed-coded since these items' framing was in the opposite directions.

Missing values

A missing value analysis displays if missing values are completely random (MCAR) or that missing values are random (MAR). The latter means that a specific question or section is responsible for the missing values. MAR is unwanted since this potentially affects the statistical power of the data set (Field, 2017). The steps of assessing the missing value analysis were: determining if the type of missing values were ignorable; the size of missing values (<10% of missing data is treated as MCAR); the randomization of missing data (e.g., Separate Variance t-Test, Little's MCAR test, and crosstabs); and selecting substitution methods in the case of MAR (e.g., listwise deletion, pairwise deletion, mean substitution, and regression techniques) (Field, 2017). Regarding the result of the missing value analyses, except for the variable 'conversational errors,' the size of the missing data for each variable is below the threshold of <5% (Hair et al., 2019). The missing value of the variable 'conversational errors' was 44.3%. This variable was added during the analyses to control for the unexpected errors during the HCI. Little's MCAR test demonstrates that the missings in the data set are completely at random $\chi 2(319) = 269.208$; $p=.259$. The implication is that there are no additional problems detected regarding the missing values.

Confirmatory Factor analyses

Performing confirmatory factor analyses (CFA) and reliability analyses upfront mitigate the common mediation effects caused by measurement errors (Field, 2017). CFA is a multidimensional technique and will display if there is enough correlation of the latent variables (factors) and the manifest variables (items) (Field, 2017). The CFA will help improve internal consistency by matching the correct items with the factors. Regarding the output of the CFA, 999 bootstrap samples had been evaluated for the multi-item constructs: Affective Empathy, Hedonic Experience, and Customer Satisfaction. Overall, the items demonstrate a high correlation on the proposed factors with values between .8 and .9 (Table 4). There were, however, issues related to discriminant validity and model fit, as displayed in table 3.

Regarding table 3 in more detail, the model fit is assed by a Bollen-Stine bootstrapping. The SMR (standardized root mean square residual) scores and the dULS (the unweighted least squares discrepancy) indicate a positive theoretical fit. However, the dG (geodesic discrepancy) score suggests otherwise, indicating an inconsistency in evaluating the model fit. For PLS path modeling, the only approximate model fit criterion is the SRMR measurement. Values below .05 indicate a good fit with a cut-off point of .08. The .0678 score on SMRS is therefore sufficient but does not demonstrates a perfect fit. The reliability of the construct scores is sufficient on the three reliability measures by surpassing the .7 threshold of the measurement criterion. The extracted variance of the factors (AVE) was above the threshold of 0.5, which indicates that each factor individually is a unidimensional construct. The HTMT estimates the (upper bound) correlation between factors. The correlations between all factors are below the threshold of <1. However, the correlation between Affective empathy and Hedonic experience was quite high .8814. The Fornell-Lackers criterion depicts: *'factor's AVE should be higher than its squared correlations with all other factors in the model.'* The squared correlation between the mentioned factors was higher than the AVE, suggesting some discriminant validity problems. Initially, three items had higher cross-loadings on theoretically unrelated factors than factors they were supposed to be loading on. These respectively were: Aff_Emp_1 on Hedonic_Exp; Aff_Emp_4 on CS, and ; Hedonic_4 on CS. The exclusion of the mentioned items increased the goodness of fit significantly. Overall, there was a sufficient approximate model fit with high loadings on the theoretical supported constructs.

Table 3

Threshold confirmatory factor analyses

Measure	**Threshold**	**Results**
Approximate model fit (saturated model)	SRMR < 0.08	0.0728
Confirmatory factor analysis (saturated model)	- SRMR < 95% bootstrap quantile (HI95 of SRMR) - dULS < 95% bootstrap quantile (HI95 of dULS) - dG < 95% bootstrap quantile (HI95 of dG)	- SRMR: sufficient - dULS: sufficient -dG: insufficient:
Internal consistency reliability	- Dijkstra-Henseler's ρA > 0.7 - Dillon-Goldstein's ρc > 0.7 - Cronbach's α > 0.7	- ρA: All constructs >0.84 - ρc: All constructs >0.83 - α: All constructs >0.82
Convergent validity	AVE > 0.5	All constructs >0.63
Discriminant validity	- HTMT significantly <1 - Fornell-Larcker criterion - Loadings exceed cross-loadings	- HTMT: all constructs <1 - Fornell-Lacker: Hed_Exp and Aff_Emp have high correlations - Three items had loadings that exceed the cross-loadings

Table 4

Item loadings

Items	Perceived Affective Empathy	Hedonic Experience	Customer satisfaction
Aff_Emp_2	0.85		
Aff_Emp_3	0.81		
Aff_Emp_5	0.83		
Aff_Emp_6	0.68		
Hedonic_1		0.82	
Hedonic_2		0.90	
Hedonic_3		0.71	
Hedonic_5		0.84	
Hedonic_6		0.84	
CS_1			0.69
CS_NPS1			0.81
CS_NPS2			0.87

Reliability analyses

The reliability analysis determines the degree of internal consistency (Cronbach's Alpha) and assesses if the elimination of items increases Cronbach's alpha level by at least > 0.05 (Field, 2017). Regarding the output, the scale to measure perceived affective empathy demonstrates a high Cronbach's Alpha = .863. By deleting the reverse coded item 'Aff_Emp_6,' Cronbach's Alpha would increase to .891. However, the increase of Cronbach's Alpha does not overweigh the lost meaning by excluding this variable. The items of hedonic experience have a high Cronbach's Alpha= .910. The deletion of items will not improve Cronbach's Alpha significantly (.014). Moreover, there is no need to delete the items related to customer satisfaction since this concept has an acceptable Cronbach's Alpha = .751. Cronbach's Alpha will not improve by deleting an item. Lastly, the total scores of the items on a construct are summed and computed on average for each respondent. This procedure is typical for regression analyses (Field, 2017).

3.4 Research ethics

The Netherlands Code of Conduct for Research Integrity principles shall be honored and regard: honesty, scrupulousness, transparency, independence, and responsibility (KNAW et al., 2018).

Honesty, by stating accurate representations of previous research and expressing the grey area of this topic like a different point of view such as the literature review CASA and the UVM theory. Moreover, the ideas and findings of different authors in the literature are not misused or refrained in the wrong context. Lastly, the current methodology of chatbot studies is not flawless and rigorous. Therefore there is a significant portion attributed to the limitations of the used scales.

Scrupulousness is established by conducting and reporting the research on the norms and rules of the scientific community. This regards authorship, such as referencing according to the official American Psychology Association (APA) reference style, referring to the source, and acknowledging authorship. Moreover, the research design includes a statistical mediation model and its corresponding analyses procedures verified in the academic context. In a more general sense, the accuracy of documentation and the quality of data collection of the research proposal needs to be verified by the assigned thesis supervisor.

Transparency is assured. Firstly, the research participants are guaranteed confidentially and anonymity. The participants have to agree to the conditions of sharing their data before conducting the survey. Secondly, the way the data is processed, used, and interpreted via statistical test to arrive at conclusions are displayed step by step. Finally, there is transparency by reporting the step-by-step approach in the methodology section.

Independence, by performing this research solely for the research institution Radboud University. Hence, there is no conflict of interest with third partners.

Responsibility had been taken into account as well. I am studied in business, not in the field of computer science and psychology. Therefore I performed a systematic literature review. The SLR helped me expand my knowledge in these fields to fully understand the essential concepts in this research. Furthermore, this study is scientific and socially relevant. Namely, chatbots and robots are taking and will be taking a more dominant role in society. Also, because these robots will become more intelligent, the role of social interaction between humans and chatbots becomes more relevant. This research addresses this interaction by exploring to which degree affection of chatbots evokes affective empathy of humans and to which degree humans attribute meaning to the expressed affection.

3.5 Methodology section of the systematic literature review

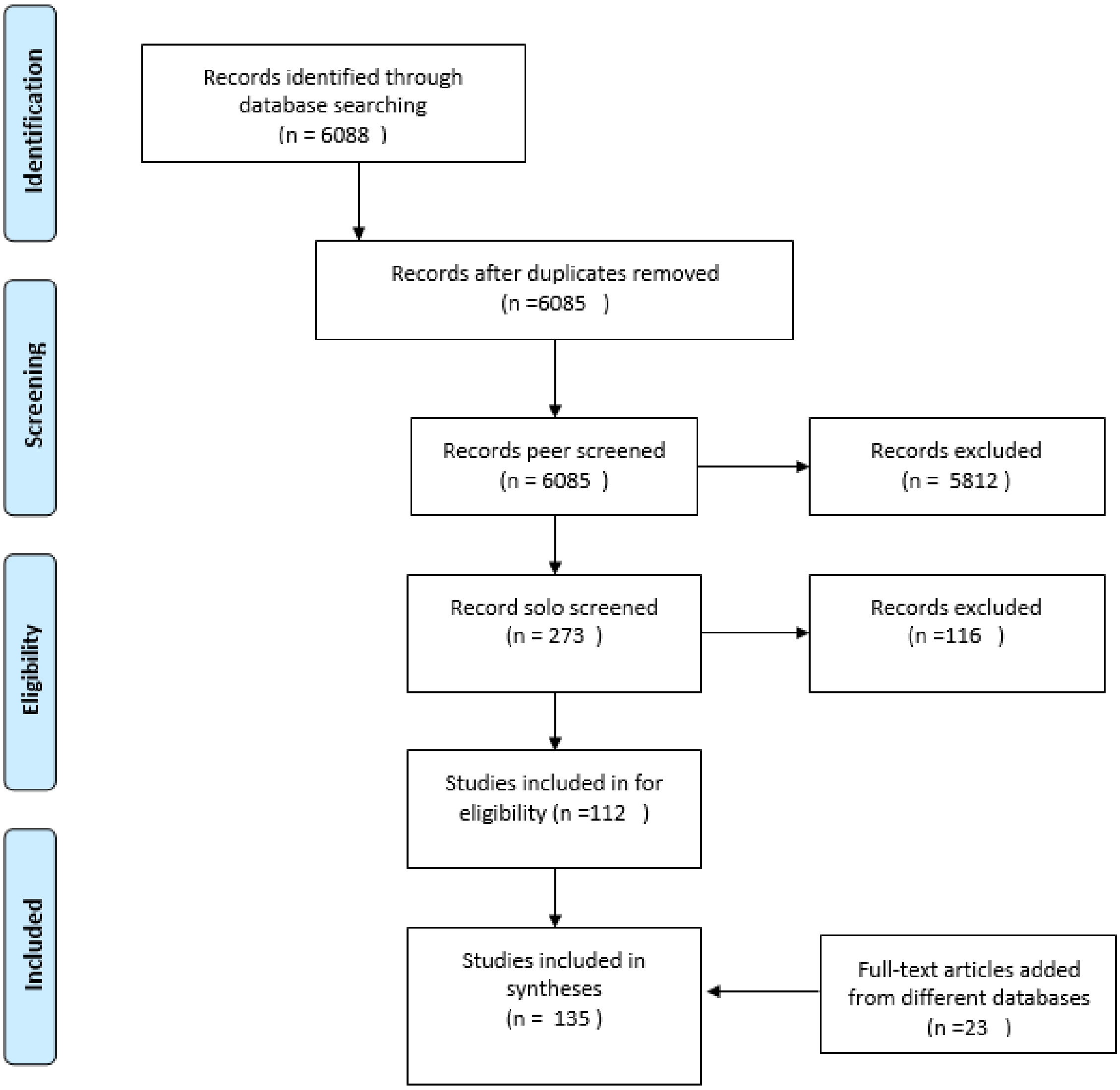
The steps of the SLR structured and framed the literature in a coherent, understandable, and verifiable way. The SLR had been peer-reviewed by MBA students who pursue similar research goals. The methodology section of the SLR regards the following phases: Plan review; Conduct review, and Document review.

Phase 1: Plan Review

To align the goals of the SLR participants, everyone shared their leading questions, together with relevant keywords, to understand each research objective. Keywords relevant to this thesis are *Customer experience, empathy, Front-end chatbots, service industries, cognitive empathy, affective empathy, HCI*. Hereafter, the group developed a review protocol that contained the inclusion and exclusion criteria of each participant. This protocol enabled the participants to better understand each work before conducting the review. Simultaneously our group reached an agreement on the appropriate search engine and the search term. The search engine is Scopus. This search engine provides peer-reviewed papers in the relevant subject field of social science and presumably contains the world's largest abstract peer-reviewed database (Enago, 2021). To ensure the quality of this search term (figure 6), we verified that the first two pages (approximately 60 of the 6088 articles) generated relevant articles. After the approval of our thesis supervisor on the search term, we developed an action plan for conducting the review.

The Scopus file was distributed to Excel and divided into 3044 articles for group A and group B, consisting of two members. Group A was to review the first 3044 articles. They had to individually assess if the articles were relevant for themselves, for another peer reviewer, or irrelevant for everyone. This assessment was done based on the inclusion and exclusion criteria of each person. After this individual assessment, the groups solved any misalignment between their codes, which ensured intercoder reliability. After clearing the misalignment within the group, both group files were combined and checked for mutual agreement of the coding.

Figure 4
Prisma flow diagram and search term



Search term

(TITLE-ABS-KEY (Empath*) AND TITLE-ABS-KEY (customer* OR experienc* OR consumer*)) AND (LIMIT-TO (LANGUAGE,"English")) AND (LIMIT-TO (EXACTKEYWORD,"Empathy"))

Phase 2 conducting the review

First, a format was designed in Excel to conduct the review in the same way. After the alignment on the format, relevant data such as the author, year, title, and abstract were extracted from Scopus and exported to Excel. The Excel metafile had been checked for duplications before both groups performed the review protocol as stated above. After peer-reviewing the articles and reaching a conclusion on the sufficiency of these articles, each researcher continued the systematic literature review on their own from that point forward. The articles which were relevant for this research were categorized into different themes based on their subject. Examples are: ‘Empathy in the service industry’ and ‘Empathy and robots.’ Henceforth, the articles were labeled on a scale from A (Very important) to D (Not important) based on the abstract and a quick scan of the article. During this phase, the conclusion was that the relation between empathy and customer experience must be explored more thoroughly. This conclusion led to the addition of more articles outside the search term (figure 4).

Phase 3: Document review

This phase contains thoroughly analyzing the involved articles of the SLR (figure 4), uncovering connections and contradictions in the literature, and writing down the findings in the final report. This phase included many iterations in which the research question changed frequently. For example, the addition of customer satisfaction as an outcome variable required more research, as was proved by the current SLR.

4. Results and Discussions

4.1 Results

Perceived affective empathy (H1a)

Perceived affective empathy demonstrates a medium to large effect size, $R^2=.50$, $F(1,93)= 94.65$, $p <.0001$. The variable which significantly predicted affective expressions is perceived affective empathy $b= 1.44$, $t(92) = 9.73$, $p<.0001$. Thus, high-level affective expressions of a chatbot lead to higher perceived affective empathy scores. For this reason, the data supports H1a.

Hedonic experience (H1b)

The effect size of hedonic experience is considerable, $R^2=.67$, $F(2,92)= 95.78$, $p <.0001$. Hedonic experience is significantly predicted by affective expressions $b= .59$, $t(92) = 3.01$, $p<.005$. Therefore, the data supports H1b, which states that perceived affective empathy increases the hedonic experience. Moreover, perceived affective empathy significantly predicts hedonic experience $b= .59$, $t(92) = 5.98$, $p<.0001$.

Customer satisfaction (H1c)

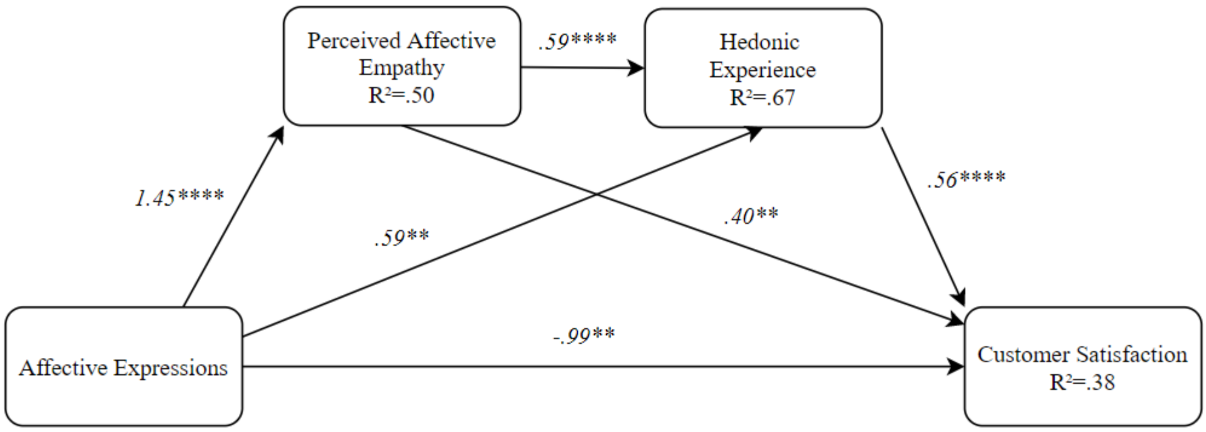
Customer satisfaction is characterized by a medium effect size $R^2=.38$, $F(3,91)= 17.60$, $p <.0001$. Hedonic experience significantly predicts customer satisfaction $b= .56$, $t(91) = 4.16$, $p<.001$. H1c is therefore supported, which states that an increase in hedonic experiences increases customer satisfaction in turn. In addition, both affective expressions, $b= -.99$, $t(91) = -3.49$, $p<.001$ as perceived affective empathy, $b= .40$, $t(91) = 2.9876$, $p<0.001$ significantly predicts customer satisfaction.

(In)direct effects on customer satisfaction (H2)

This research expects that the long-way specific indirect effect of affective expressions on customer satisfaction (Ind3) best explains the relationship between X and Y than the direct effect and the short cut specific indirect effects (Ind1, Ind2). The results indicate that perceived affective empathy and hedonic experience mediated the relationship between affective expressions and customer satisfaction, Ind3=.48, SE= .16, 95% [0.21, 0.82]. Perceived affective empathy mediates the relationship between affective expressions and customer satisfaction, ind1=.57, SE= .20, 95% [0.15, 0.96]. Hedonic experience mediates the relationship between affective expressions and customer satisfaction, ind2 =.33, SE= .13, 95% [0.21, 0.82]. The direct effect of affective expressions on customer satisfaction is significant and negatively related to customer satisfaction b= -.99, t(91) = -3.49, p<.001. The data supports H2b since the effect size of Ind3 outweigh Ind2. We must reject H2a since the effect size of Ind1 is larger than the effect size of Ind3. Due to the unexpected stronger and opposite direct effect compared to Ind3, we must reject H2c.

Figure 5

Results: Serial mediation model



*Symbol Meaning: Ns = P > 0.05; * = P ≤ 0.05; **= P ≤ 0.01; ***= P ≤ 0.001; ****= P ≤ 0.0001*

Table 5

Output serial mediation without covariates

	B	SE (HC0)	t	p
Predictors				
Outcome: Perceived Affective Empathy (R^2 = .50)				
Affection *(H1a)*	**1.45**	.15	9.73	.000
Outcome: Hedonic Experience (R^2 = .67)				
Affection	**.59**	.20	3.01	.003
Perceived Affective Empathy *(H1b)*	**.59**	.10	5.97	.000
Outcome: Customer Satisfaction (R^2 = .38)				
Affection	**(-).99**	.28	(-)3.49	.007
Perceived Affective Empathy	**.40**	.13	2.97	.004
Hedonic Experience *(H1c)*	**.56**	.14	4.16	.0001

Indirect effect on customer satisfaction	B	BootSE	95% CI
Ind1 (x->m1->y)	**.57**	.20	[0.15, 0.96]
Ind2 (x->m2->y)	**.33**	.13	[0.11, 0.62]
Ind3 (x->m1->m2->y)	**.48**	.16	[0.21, 0.82]

4.2 Post hoc analyses

This study supports the long-way specific indirect effect of affective expressions, perceived affective empathy, and hedonic experience on customer satisfaction. However, this model is a competitive partial mediation model due to the significance of both casual indirect as direct effects with opposing signs (L. Zhou et al., 2010). The implication is that there were omitted mediators moderators left out of the analyses. The absence causes a risk of underestimating the importance of the mediation process (Shrout & Bolger, 2002). The post hoc analysis unravels the issues mentioned above.

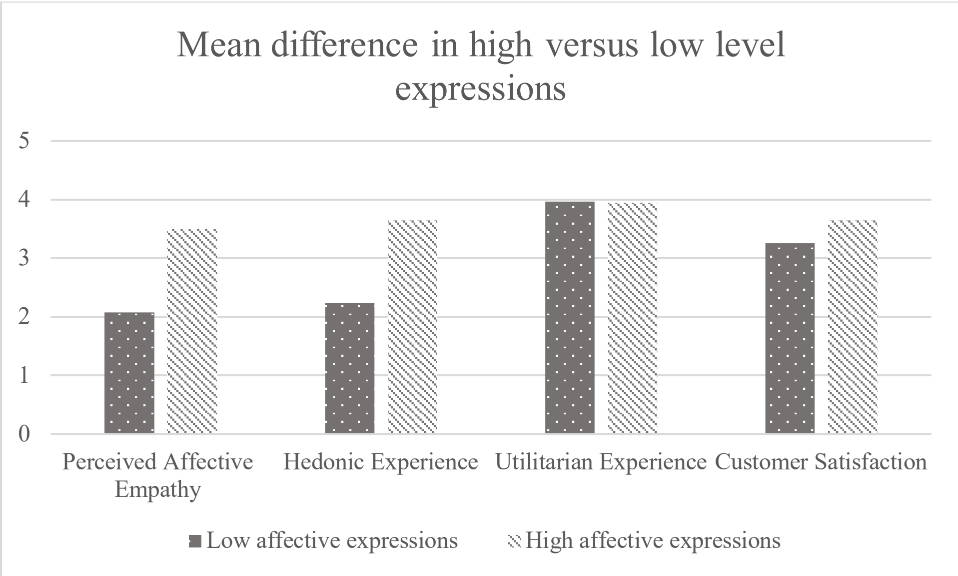
Mean differences between low-level and high-level affective expressions

The overall level of satisfaction of high-level chatbot expressions (M= 3.65, SD=.15) was higher than the low level of chatbot expressions (M=.3.25, SD=.15). Nonetheless is the differential margin on the overall satisfaction scores between both levels relatively small. Therefore, the respondents were reasonably satisfied with both chatbots. The chatbot with high-level expressions indicates relatively high hedonic experience scores (M=.3.65, SD=.79), whereas the low-level affective chatbot (M=.2.24, SD=.70) creates relatively low hedonic experiences. The explanation for the low hedonic experience might be related to this chatbot's perceived affective empathy score, M=2.07, which was higher for the high-level affective chatbot, M=3.50.

High scores on perceived affective empathy and hedonic experience do not lead to substantially higher customer satisfaction than low scores on these variables. The implication is that perceived affective empathy and hedonic experience does not fully explain customer satisfaction. A simple linear regression indicates that utilitarian experience significantly predicts customer satisfaction, $b = .922$, $t(92) = 12.30$, $p < .001$. For this reason, the utilitarian experience of the high-level affective chatbot (M=.3.94, SD=.87) and the low-level affective chatbot (M=.3.93, SD=.95) demonstrates a generally useful experience which caused the respondents to be satisfied.

Figure 6

Mean differences in high-level versus low-level affective expressions



Unraveling the direct adverse effect

As mentioned, an omitted variable causes the negative direct effect of affective expressions on customer satisfaction (Figure 5). A simple linear regression of affective expressions on customer satisfaction, R^2=.03, $F(1,93)$= 3.40, p =.07, is insignificant. Hence, more information is needed to explain the relationship between affective expressions and customer satisfaction to understand this (direct) adverse effect. The expectation is that utilitarian experience is this omitted variable. Namely, utilitarian experience predicts customer satisfaction, the chatbots both have similar high utilitarian scores, but dissimilar hedonic experience and perceived affective empathy scores, and similar customer satisfaction scores. In continuation of this paragraph, a series of analyses demonstrate the exact impact of utilitarian experience on the serial mediation model. Table 6 contains a summary of these analyses.

Firstly, hedonic experience (x1) and utilitarian experience (x2) significantly predict customer satisfaction (y), R^2=.69, $F(2,92)$= 103.40, p <.001. Both utilitarian experience, b= .86, t(94) = -11.32, p<.001 as hedonic experience, b= .25, t(94) = -3.81, p<..001 lead to a positive customer satisfaction. However, the effect of the standardized coefficients=.71 of utilitarian experience is more impactful than the standardized coefficients=.24 of hedonic experience.

Secondly, by taking into account affective expressions (x) and utilitarian experience (y) in a simple mediation model, it comes to the surface that utilitarian experience (y),

R^2=.001, F(1,93)= .009, p =.76. is not significantly predicted by affective expressions. Customer satisfaction (y), R^2=.67, F(2,92)= 119.07, p <.001 is predicted by utilitarian experience (mediator), b= .96, t(92) = 15.26, p<.001 and directly through affective expressions (x), b= .34, t(94) = 2.78, p <.01. These results imply and confirm that both chatbots offer a similar utilitarian experience, which improves customer satisfaction. Although more expressed affection directly leads to higher customer satisfaction scores.

Thirdly, by performing a simple mediation analysis with hedonic experience instead of utilitarian experience, we see that just as in figure 5, hedonic experience (y), R^2=.52, F(1,93)= 104.99, p <.001 is predicted by affective expressions (x), b= 1.48, t(93) = 10.25, p <.001. Customer satisfaction (y), R^2=.34, F(2,92)= 21.60, p <.001 is predicted by hedonic experience (mediator), b= .82, t(93) = 5.72, p <.001 and affective expression (x), b= -.82, t(93) = -2.76, p <.01. Therefore as expected, high-level affective expression (Esmee) leads to more perceived affective empathy, increasing customer satisfaction consequently. The direct effect is negative, indicating that low-level affective expressions (Cassandra) directly increases customer satisfaction. The relatively low explained variance, R^2=.34 of customer satisfaction, underlines some unexplained variance in predicting customer satisfaction. The previous two post hoc analyses, which included utilitarian experience as a predictor of customer satisfaction, better explain the variance of customer satisfaction (R^2=.69 and R^2=.67) compared to the models which excluded this variable.

Lastly, utilitarian experience (y), R^2=.30, F(2,92)= 12.58, p <.001 had successfully been predicted by affective expressions (x), b= -.92, t(92) = -3.75, p <.001 and perceived affective empathy (mediator), b= .67, t(92) = 5.00, p <.001 in a simple mediation model. The result indicates that low-level affective expressions directly lead to a more utilitarian experience. High-level affective expressions are positively related to perceived affective empathy, leading to a better utilitarian experience.

To summarize these findings, both chatbots offer relatively high utilitarian experiences. Moreover, the utilitarian experience had a more substantial impact on customer satisfaction compared to the hedonic counterpart. Furthermore, utilitarian experience explains a high degree of customer satisfaction variance. Ultimately, low-level affective expression directly leads to utilitarian experience, whereas high-level affective expression leads to a utilitarian experience directly and indirectly through perceived affective. These findings conclude that the absence of utilitarian experience in the proposed model causes the negative direct effect of affective expressions on customer satisfaction (Figure 5). With the inclusion of utilitarian experience in this model, the negative direct effect of affective expressions on customer satisfaction will become insignificant, and the low-level affective expressions will lead to customer satisfaction via utilitarian experience.

Table 6

Summary post hoc tests 'Unraveling the direct adverse effect.'

Analyses	Dependent variable	Predictor 1	Predictor 2	Variance Explained
1	CS	Uti_Exp, (x1), b= .86	Hed_Exp, (x2), b= .25	R^2=.69 ***
2.1	Uti_Exp	Aff_Expre, = NS.		R^2=NS
2.2	CS	Aff_Expre (x), b= .34	Uti_Exp (m), b= .96	R^2=.67***
3.1	Hed_Exp	Aff_Expre, b= 1.48		R^2=.52***
3.2	CS	Aff_Expre (x), b= -.82	Hed_Exp (m), b= .82	R^2=.34***
4	Uti_Exp	Aff_Expre (x), b= -.92	Per_Aff_Emp (m), b= .67	R^2=.30***

Note. *CS = Customer satisfaction, Uti_Exp = Utilitarian Experience, Hed_Exp = Hedonic experience, Aff_Expre = Affective Expressions, Per_Aff_Emp = Perceived Affective Empathy*

Effect of the covariates

The covariates do not significantly predict perceived affective empathy and hedonic experiences (Table 7). The covariate conversation errors significantly predicts customer satisfaction b= -.73, t(40)= -5.93, p<001. The remaining covariates do not significantly predict customer satisfaction (Table 7). Perhaps does the conversational error further explains the difference in customer satisfaction scores. As beforementioned this variable is a significant predictor of customer satisfaction, b= -.73, t(40)= -5.93, p<001. However, the difference between the two chatbots for the amount of conversational error did not differ, $\chi2(2, N = 54) = 0.15$, p = .93. Meaning that conversational errors did not significantly impact customer satisfaction more in favor of one chatbot above the other.

Learning process

Between the period '18 June - 1 July', the respondents conversed with the chatbots 'Cassandra' and 'Esmee.' Google Dialogue flow allowed to monitor these conversations and correct mistakes of the chatbots if these occurred. Since conversational error negatively influences customer satisfaction, and the learning process of the chatbot aims to reduce conversational error, satisfaction scores over time could increase and bias this research. Fortunately, the learning process of the chatbots did not affect customer satisfaction, R^2=.001, $F(1,93)$= .108, p =.744, which means that there was no significant improvement in customer satisfaction over time.

Table 7

Output serial mediation analyses with covariates

	B	SE (HC0)	t	p
Predictors				
Outcome: Perceived Affective Empathy (R^2 = .57)				
Affection *(H1a)*	**1.38**	.20	6.85	.00
Past interaction	.25	.20	1.29	.20
Attitude	.14	.11	1.32	.19
Susc.empathy	.15	.09	1.60	.11
Gender	(-).06	.19	(-).33	.74
Conversational errors	(-).28	.14	(-)1.96	.06
Outcome: Hedonic Experience (R^2 = .78)				
Affection	**.63**	.18	3.5	.001
Perceived Affective Empathy *(H1b)*	**.56**	.08	6.82	.000
Past interaction	.10	.14	.66	.51
Attitude	.09	.07	1.22	.23
Susc.empathy	.01	.06	.19	.85
Gender	.15	.13	1.10	.28
Conversational errors	(-).05	.11	(-).50	.62
Outcome: Customer Satisfaction (R^2 = .66)				
Affection	**(-).78**	.34	(-)2.31	.004
Perceived Affective Empathy	(-).15	.19	(-).76	.4492
Hedonic Experience *(H1c)*	**.89**	.17	5.26	.000
Past interaction	(-).06	.19	(-).33	.74
Attitude	.06	.11	.52	.61
Susc.empathy	(-).05	.09	(-).52	.60
Gender	(-).02	.17	(-).13	.90
Conversational errors	**(-).73**	.12	(-)5.93	.000
Indirect effect on customer satisfaction	B	BootSE	95% CI	
Ind1 (x->m1->y)	(-).20	.31	[-0.90, 0.38]	
Ind2 (x->m2->y)	**.52**	.28	[0.15, 1.23]	
Ind3 (x->m1->m2->y)	**.64**	.23	[0.31, 1.19]	

5. Discussion

5.1 Key contributions

This study's key contributions are (1) evidence of the proposed serial mediation effect which states that the indirect effect of high-level affective chatbot expressions creates perceived affective empathy, hedonic experience, and higher customer satisfaction scores consequently, (2) utilitarian experience explains why low-level affective chatbots could also lead to higher customer satisfaction scores, and (3) Conversational errors decrease customer satisfaction scores significantly.

Theoretical contributions

Regarding the first main contribution, (1) the R^2 of the dependent variables within the serial mediation model was robust, meaning that the data had explained much variance. Hypotheses 1 was fully supported; hence there was support for the long-way-indirect effect. The acceptance of hypothesis 1 is in line with the findings of Verhagen (2014) and Liu & Sundar (2018). They found out that by manipulating language style, service agents are perceived as more friendly by customers. Previous research on service settings found that empathy leads to higher satisfaction in a positive service setting compared to a negative service setting (C. Davis et al., 2017). However, this research, placed in a negative service setting, did indicate that an increase in perceived affective empathy increases customer satisfaction scores. Furthermore, the expectation was that the long-way-indirect effect (Ind3) was the most dominant (indirect)effect. However, ind1 had a slightly stronger beta compared to ind3. Thus, the high-level affective expressions of a chatbot leading to perceived affective empathy alone are enough to create higher customer satisfaction scores. Nonetheless, the addition of hedonic experience through Ind3 explains more variance and nuances in how customer satisfaction is affected through high-level affective expressions. The research findings contribute to the call of Gentile et al. (2007) to explore intangible elements such as customers' emotions in the field of experiential marketing. Moreover, by confirming the role of affection in creating a hedonic experience through a chatbot interaction, this research contributed to research on customer happiness and the quality of life in creating meaningful experiences (Jain et al., 2017). This research advocates the CASA theory over the UVM theory since no uncanny attitude of the participants had been reflected in this research. Uncanniness would have otherwise been spotted in low perceived scores of affective empathy, customer experience, and customer satisfaction score of Esmee, the high-level affective expression chatbot.

Concerning the second main contribution (2), an unexpected but essential observation was the direct adverse effect of affective expressions on customer satisfaction. In the post hoc analyses, the absence of utilitarian experience in the model explained this effect. The strong impact of utilitarian experience relates perhaps to the context 'damage claim at an insurance agency. Khedhaouria and Beldi's (2014) research on customer banking and text messaging and Ryu et al. (2010) on fast-food restaurants had a similar finding in which utilitarian experience outweighs the importance of hedonic experience. Furthermore, it contradicts the findings of shopping experiences (Babin et al., 1994; Eroglu et al., 2005). Another explanation of the strong impact of utilitarian experience relates to the motivations of using a chatbot: convenience and efficiency (Brandtzaeg & Følstad, 2017; Rzepka et al., 2020).

The third main contribution (3), the conversational errors were present in both chatbots. Even by controlling for the scenario and creating conversational guidelines for the participants, conversational errors were not excluded. This finding underlines the complexity of designing HCI, where interaction could lead to complex conversations that ended up being misinterpreted or misunderstood (Kocaballi et al., 2020). Therefore, his research supports the call of Balkenius et al. (2016) on autonomy, capabilities, and the intrinsic morality of chatbots in creating human-like conversations.

Managerial contributions

Pertain to the essential contributions; this study states that high-level affective expressions as low-level affective expressions evoke different perceived affective empathic levels, different experiences but marginal similar customer satisfaction scores. The implication for the practical field is that chatbot developer should create a utilitarian experience at a minimum level and reduce conversational errors. However, besides brands associated with hedonic experiences such as shopping experience (Babin et al., 1994; Eroglu et al., 2005), this research demonstrates that the service industry, specifically the insurance sector, can use high-level affective expressions of chatbots in their service offering. Moreover, expressed high-level affection of chatbots can deepen the relationship between companies and customers by evoking perceived affective empathy of customers. Designing a more empathic chatbot gives more insight into the soft knowledge of customers, such as human cognition and emotions. Soft knowledge helps design robots that behave according to appropriate social behavior (Broadbent, 2017). This knowledge is imperative to understand and implement virtual agents' design technology that enables knowledge sharing to create meaningful online experiences (Z. Zhou et al., 2012). There is always a risk of too emotionally charged messages, which leads to perceived eeriness (Stein & Ohler, 2017). Besides, the effect of expressed empathy on attitudes could occur weeks after the displayed interaction (Batson et al., 1997). However, linguistics such as text messages could be interpreted accordingly and matched with the proper response to align to the individual's affective state (Balkenius et al., 2016).

5.2 Limitations and future work

Limitations

There were a few limitations detected related to the experimental setting. First of all, this research did not control the participants' emotional state before the conversation. Therefore it cannot be ruled out that external factors influence the mood of the participants leading to different results. Moreover, the chatbot interaction took place in a demo version of Google DialogFlow. This environment had limitations of its own. Firstly, there was a fixed text in the message box with the words ‘ask something.’ This fixed message contradicts the instruction, which was to greet the chatbot. Secondly, there was a fixed microphone symbol in the message box, which suggested the participants use it. These confusions were reduced by plurally mentioning the instructions in the survey. However, it might have occurred that participants tried and use this microphone. Thirdly, as controlled for, conversational errors affected customer satisfaction. Clear instructions and a narrowed-down scenario to a few minutes of interaction should have significantly reduced these errors. Lastly, the chatbots were not only distinct through their expressions, but in their looks as well. The low-level affective chatbot did not have a human face as an avatar, unlike the high-level affective chatbot. There is no assessment made in this research to which degree facial avatars impacted the results. Previous research shows that artificial pictures did not decrease perceived empathy towards the digital persona (Salminen et al., 2020). However, controlling for this effect would have confirmed these findings, making the study more robust.

An additional limitation is related to the sample size. Although 189 participants opened the survey, only 98 had finished it. The drop-out is primarily related to the moment that participants had to converse with the chatbots. Although the sample size was large enough, and the serial mediation model had a high explaining power, many covariates were insignificant. Perhaps a large sample would increase the explaining power of these covariates to a more considerable degree, especially since the previous theory expects these variables to be significant. Moreover, convenience sampling is limited because non-probability sampling does not assume representativity, negatively impacting the external validity of the findings (Field, 2017).

Moreover, is caution in interpreting the scales important. THe RoPE scale is new and therefore is less robust (Charrier et al., 2019). However, this scale is an adaptation of the IRI scale. The IRI scale is a validated measurement scale of perceived affective empathy because of the multidimensional conceptualization of empathy, the comprehensive measurement of self-reported empathy, and its parsimony (De Corte et al., 2007). THe RoPE scale is used due to its accurate adaption of IRI in the HRI context and therefore offers high content validity and face validity. The low robustness of the RoPE scale does, however, asks for caution in interpreting the results of the RoPE scale. Regarding the HED/UT scale, Heijden & Sørensen (2003) suggested that more empirical evidence through independent and representative samples is needed to gain complete confidence in the scale. The researchers, therefore, suggest being careful in the interpretation of this scale (Heijden & Sørensen, 2003).

Furthermore, one of the disadvantages of PROCESS was that the estimation procedures are based on the observed variables. By not including latent variable measurement models as is done with structural equation modeling (SEM), the measurement error in the estimation process is less profound (Hayes, 2012).

Lastly, there was considerable overlap between the items of hedonic experience and perceived affective empathy. For this reason, items with cross-loadings had been excluded from the analyses to reduce discriminant validity. This exclusion led to a more distinct separation between these concepts but also limited the information available to interpret these two concepts.

Future work

Firstly, it is vital to assess to which degree the main results of this research hold up between industries since previous research report different valuation of hedonic versus utilitarian experience between industries (Babin et al., 1994; Eroglu et al., 2005; Khedhaouria

& Beldi, 2014; Ryu et al., 2010). In addition, because cultural aspects such as language differences vary among geographical regions, it is valuable to assess country-specific valuation on affective expressions of chatbots (Hill et al., 2015).

Secondly, the assessment of this research is as a longitudinal study shall provide meaningful results. Since perceived empathy could be felt days or even weeks after the interaction (Batson et al., 1997), there is more nuance on the perceived affective empathy levels over time and its overall impact on the proposed model (Figure 5). This longitudinal study aligns with the call for more research on long-term Human-Robot-Interaction (de Gennaro et al., 2020; Paiva et al., 2017). Customer satisfaction of the current study regards transaction-specific satisfaction. Longitudinal research could measure cumulative satisfaction, meaning customer satisfaction over time (Johnson et al., 1995), which richen our understanding of long-term HCI. Thirdly, more manipulations of the chatbot enlarge our knowledge of the effect of chatbot affection. We already see a focus on non-verbal communication elements such as facial expressions in the gaming industry (Ho & Ng, 2020; Prendinger et al., 2006; Sierra Rativa et al., 2020; Swiderska & Küster, 2018; Wilde & Evans, 2019). A more extreme manipulation regards the comparison of different digital entities (e.g., virtual humans, voice assistants, human customer service) on textual versus verbal affection expressed by these entities.

5.2 Conclusion

Chatbots in the service sector are associated with practical experiences. This association underlines the current undervaluation of empathy in the service sector. The findings of this study indicate that chatbots do not necessarily offer utilitarian experiences alone but can offer hedonic experiences through high-level affective expressions. The results support the primary research proposition that high-level affective expressions lead to higher customer satisfaction scores. Perceived affective empathy and hedonic experience respectively mediate this effect. Moreover, this research underlines that the low-level affective expressions chatbot create solely utilitarian experiences and high-level affective expressions chatbot, on the other hand, creates both utilitarian experiences as hedonic experiences. Lastly, although high-level expressed affection led to higher customer satisfaction scores, the utilitarian experience is relatively more critical than the hedonic experience for customers to be satisfied.

References

Abubshait, A., Beatty, P. J., McDonald, C. G., Hassall, C. D., Krigolson, O. E., & Wiese, E. (2021). A win-win situation: Does familiarity with a social robot modulate feedback monitoring and learning? *Cognitive, Affective and Behavioral Neuroscience*, 1–13. https://doi.org/10.3758/s13415-021-00895-9

ADANCO. (n.d.). *Equation Modeling - ADANCO by Composite Modeling*. Retrieved July 6, 2021, from https://www.composite-modeling.com/

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Araújo, T., & Casais, B. (2020). Customer Acceptance of Shopping-Assistant Chatbots. *Smart Innovation, Systems and Technologies*, *167*, 278–287. https://doi.org/10.1007/978-981-15-1564-4_26

Ayanouz, S., Abdelhakim, B. A., & Benhmed, M. (2020). A Smart Chatbot Architecture based NLP and Machine Learning for Health Care Assistance. *ACM International Conference Proceeding Series*, 1–6. https://doi.org/10.1145/3386723.3387897

Babin, B. J., Darden, W. R., & Griffin, M. (1994). Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value. *Journal of Consumer Research*, *20*(4), 644. https://doi.org/10.1086/209376

Bachelor, A. (1988). How clients perceive therapist empathy: A content analysis of “received” empathy. *Psychotherapy*, *25*(2), 227–240. https://doi.org/10.1037/h0085337

Bagozzi, R. P., & Burnkrant, R. E. (1979). Attitude organization and the attitude-behavior relationship. *Journal of Personality and Social Psychology*, *37*(6), 913–929. https://doi.org/10.1037/0022-3514.37.6.913

Balkenius, C., Cañamero, L., Pärnamets, P., Johansson, B., Butz, M. V., & Olsson, A. (2016). Outline of a sensory-motor perspective on intrinsically moral agents. *Adaptive Behavior*, *24*(5), 306–319. https://doi.org/10.1177/1059712316667203

Barnes, D. C., Ponder, N., & Dugar, K. (2011). Investigating the key routes to customer delight. *Journal of Marketing Theory and Practice*, *19*(4), 359–375. https://doi.org/10.2753/MTP1069-6679190401

Bartneck, C., & Reichenbach, J. (2005). Subtle emotional expressions of synthetic characters. *International Journal of Human Computer Studies*, *62*(2), 179–192. https://doi.org/10.1016/j.ijhcs.2004.11.006

Batra, R., & Ahtola, O. T. (1990). Measuring the Hedonic and Utilitarian Sources of Consumer Attitudes. In *Marketing Letters* (Vol. 2, Issue 2).

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and Empathy: Two Qualitatively Distinct Vicarious Emotions with Different Motivational Consequences. *Journal of Personality*, *55*(1), 19–39. https://doi.org/10.1111/j.1467-6494.1987.tb00426.x

Batson, C. D., Klein, T. R., Highberger, L., & Shaw, L. L. (1995). Immorality from empathy-induced altruism: When compassion and justice conflict. *Journal of Personality and*

Social Psychology, *68*(6), 1042–1054. https://doi.org/10.1037//0022-3514.68.6.1042

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, *72*(6), 1335–1348. https://doi.org/10.1037/0022-3514.72.6.1335

Batson, C. D., & Shaw, L. L. (1991). Evidence for Altruism: Toward a Pluralism of Prosocial Motives. *Psychological Inquiry*, *2*(2), 107–122. https://doi.org/10.1207/s15327965pli0202_1

Batson, C. D., Turk, C. L., Shaw, L. L., & Klein, T. R. (1995). Information Function of Empathic Emotion: Learning That We Value the Other's Welfare. *Journal of Personality and Social Psychology*, *68*(2), 300–313. https://doi.org/10.1037/0022-3514.68.2.300

Bendapudi, N., Singh, S. N., & Bendapudi, V. (1996). Enhancing Helping Behavior: An Integrative Framework for Promotion Planning. *Journal of Marketing*, *60*(3), 33–49. https://doi.org/10.1177/002224299606000303

Bendle, N. T., Farris, P. W., Pfeifer, P. E., & Reibstein, D. J. (2016). *MARKETING METRICS THIRD EDITION* (A. Neidlinger (Ed.); 3rd ed.). Boger, Paul .

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. *377-392*. https://doi.org/10.1007/978-3-319-70284-1_30

Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human Computer Studies*, *62*(2), 161–178. https://doi.org/10.1016/j.ijhcs.2004.11.002

Broadbent, E. (2017). Interactions With Robots: The Truths We Reveal About Ourselves. *Annual Review of Psychology*, *68*(1), 627–652. https://doi.org/10.1146/annurev-psych-010416-043958

Charrier, & Galdeano. (2018). *(PDF) Empathy Display Influence on Human-Robot Interactions: a Pilot Study*. https://www.researchgate.net/publication/328289664_Empathy_Display_Influence_on_Human-Robot_Interactions_a_Pilot_Study

Charrier, L., Rieger, A., Galdeano, A., Cordier, A., Lefort, M., & Hassas, S. (2019). The RoPE Scale: A Measure of How Empathic a Robot is Perceived. *ACM/IEEE International Conference on Human-Robot Interaction*, *2019-March*, 656–657. https://doi.org/10.1109/HRI.2019.8673082

Chen, N., Mohanty, S., Jiao, J., & Fan, X. (2021). To err is human: Tolerate humans instead of machines in service failure. *Journal of Retailing and Consumer Services*, *59*. https://doi.org/10.1016/j.jretconser.2020.102363

Chitturi, R., Raghunathan, R., & Mahajan, V. (2008). Delight by Design: The Role of Hedonic versus Utilitarian Benefits. *Journal of Marketing*, *72*(3), 48–63. https://doi.org/10.1509/jmkg.72.3.048

Chopra, A. (2020). *21 Vital Chatbot Statistics for 2020 | Interactive Content on Outgrow*. https://outgrow.co/blog/vital-chatbot-statistics

Chopra, K. (2019). Indian shopper motivation to use artificial intelligence: Generating Vroom's expectancy theory of motivation using grounded theory approach. *International*

Journal of Retail and Distribution Management, *47*(3), 331–347. https://doi.org/10.1108/IJRDM-11-2018-0251

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

Clark, C. M., Murfett, U. M., Rogers, P. S., & Ang, S. (2013). Is Empathy Effective for Customer Service? Evidence From Call Center Interactions. *Journal of Business and Technical Communication*, *27*(2), 123–153. https://doi.org/10.1177/1050651912468887

Collier, J. E., Barnes, D. C., Abney, A. K., & Pelletier, M. J. (2018). Idiosyncratic service experiences: When customers desire the extraordinary in a service encounter. *Journal of Business Research*, *84*, 150–161. https://doi.org/10.1016/j.jbusres.2017.11.016

Conner, M., & Armitage, C. J. (1998). Extending the Theory of Planned Behavior: A Review and Avenues for Further Research. *Journal of Applied Social Psychology*, *28*(15), 1429–1464. https://doi.org/10.1111/j.1559-1816.1998.tb01685.x

Costa, G., Glinia, E., & Drakou, A. (2004). The role of empathy in sport tourism services: A review. *Journal of Sport and Tourism*, *9*(4), 331–342. https://doi.org/10.1080/1477508052000341887

Cross, E. S., Riddoch, K. A., Pratts, J., Titone, S., Chaudhury, B., & Hortensius, R. (2019a). A neurocognitive investigation of the impact of socializing with a robot on empathy for pain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1771). https://doi.org/10.1098/rstb.2018.0034

Cross, E. S., Riddoch, K. A., Pratts, J., Titone, S., Chaudhury, B., & Hortensius, R. (2019b). A neurocognitive investigation of the impact of socializing with a robot on empathy for pain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1771). https://doi.org/10.1098/rstb.2018.0034

Davis. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126. https://doi.org/10.1037/0022-3514.44.1.113

Davis, C., Jiang, L., Williams, P., Drolet, A., & Gibbs, B. J. (2017). Predisposing Customers to Be More Satisfied by Inducing Empathy in Them. *Cornell Hospitality Quarterly*, *58*(3), 229–239. https://doi.org/10.1177/1938965517704373

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly: Management Information Systems*, *13*(3), 319–339. https://doi.org/10.2307/249008

De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., & Davis, M. H. (2007). Measuring Empathic Tendencies: Reliability And Validity of the Dutch Version of the Interpersonal Reactivity Index. *Psychologica Belgica*, *47*(4), 235. https://doi.org/10.5334/pb-47-4-235

de Gennaro, M., Krumhuber, E. G., & Lucas, G. (2020). Effectiveness of an Empathic Chatbot in Combating Adverse Effects of Social Exclusion on Mood. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.03061

Demming, C. L., Jahn, S., & Boztug, Y. (2017). Conducting Mediation Analysis in Marketing Research. *Marketing ZFP*, *39*(3), 76–98. https://doi.org/10.15358/0344-1369-2017-3-76

Dialogflow. (n.d.). Retrieved May 30, 2021, from https://dialogflow.cloud.google.com/#/getStarted

Diefenbach, S., & Hassenzahl, M. (2011). The dilemma of the hedonic - Appreciated, but hard to justify. *Interacting with Computers*, *23*(5), 461–472. https://doi.org/10.1016/j.intcom.2011.07.002

Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, *94*(1), 100–131. https://doi.org/10.1037/0033-2909.94.1.100

Eisenberg, N., & Miller, P. A. (1987). The Relation of Empathy to Prosocial and Related Behaviors. In *Psychological Bulletin* (Vol. 101, Issue 1, pp. 91–119). https://doi.org/10.1037/0033-2909.101.1.91

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Enago. (2021). *ISI, Scopus or PubMed: How Do I Choose the Best Journal Based on its Indexing? - Enago Academy*. https://www.enago.com/academy/isi-scopus-or-pubmed/

Eroglu, S. A., Machleit, K., & Barr, T. F. (2005). Perceived retail crowding and shopping satisfaction: The role of shopping values. *Journal of Business Research*, *58*(8), 1146–1153. https://doi.org/10.1016/j.jbusres.2004.01.005

Etemad-Sajadi, R. (2014). The influence of a virtual agent on web-users' desire to visit the company: The case of restaurant's web site. *International Journal of Quality and Reliability Management*, *31*(4), 419–434. https://doi.org/10.1108/IJQRM-05-2013-0077

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. . (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96–104. https://doi.org/10.1145/2818717

Feshbach, N. D. (1975). Empathy in Children: Some Theoretical and Empirical Considerations. *The Counseling Psychologist*, *5*(2), 25–30. https://doi.org/10.1177/001100007500500207

Field, A. (2017). *Discovering Statistics Using IBM SPSS* (5th ed.). Sage Publications Ltd.

Fishbein, M., & Ajzen, I. (1977). Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research - PhilPapers. *Philosophy and Rhetoric*, *10*(2). https://philpapers.org/rec/FISBAI

Gentile, C., Spiller, N., & Noci, G. (2007). How to Sustain the Customer Experience:. An Overview of Experience Components that Co-create Value With the Customer. *European Management Journal*, *25*(5), 395–410. https://doi.org/10.1016/j.emj.2007.08.005

Glaskin, K. (2012a). Empathy and the robot: A neuroanthropological analysis. *Annals of Anthropological Practice*, *36*(1), 68–87. https://doi.org/10.1111/j.2153-9588.2012.01093.x

Glaskin, K. (2012b). EMPATHY AND THE ROBOT: A NEUROANTHROPOLOGICAL ANALYSIS. *Annals of Anthropological Practice*, *36*(1), 68–87. https://doi.org/10.1111/j.2153-9588.2012.01093.x

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619. https://doi.org/10.1126/science.1134475

Hayes, A. F. (2012). *PROCESS: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling 1*. http://www.afhayes.com/

Heijden, H., & Sørensen, L. S. (2003). Measuring attitudes towards mobile information services: an empirical validation of the HED/UT scale. *Undefined*.

Helpshift. (2019). *State of Customer Service Automation 2019 Insights and trends based on analysis of 75 million customer service tickets and 71 million bot interactions*. https://go.helpshift.com/rs/113-UDX-599/images/Report_State_of_CS_Automation.pdf

Hill, J., Randolph Ford, W., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior*, *49*, 245–250. https://doi.org/10.1016/j.chb.2015.02.026

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic Consumption: Emerging Concepts, Methods. In *Source: Journal of Marketing* (Vol. 46, Issue 3).

Ho, J. C. F., & Ng, R. (2020). Perspective-Taking of Non-Player Characters in Prosocial Virtual Reality Games: Effects on Closeness, Empathy, and Game Immersion. *Behaviour and Information Technology*. https://doi.org/10.1080/0144929X.2020.1864018

Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The Influence of Avatars on Online Consumer Shopping Behavior. *Journal of Marketing*, *70*(4), 19–36. https://doi.org/10.1509/jmkg.70.4.019

Hortensius, R., & Cross, E. S. (2018). From automata to animate beings: the scope and limits of attributing socialness to artificial agents. *Annals of the New York Academy of Sciences*, *1426*(1), 93–110. https://doi.org/10.1111/nyas.13727

Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology*, *22*(4), 582–594. https://doi.org/10.1016/j.jcps.2012.03.006

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology*, *17*(2), 139–153. https://doi.org/10.1016/S1057-7408(07)70020-7

IBM SPSS Statistics 26. (n.d.). Retrieved May 30, 2021, from https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-26

Isbister, K., & Nass, C. (2000). Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *International Journal of Human Computer Studies*, *53*(2), 251–267. https://doi.org/10.1006/ijhc.2000.0368

Ischen, C., Araujo, T., Voorveld, H., van Noort, G., & Smit, E. (2020). Privacy Concerns in Chatbot Interactions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11970 LNCS*, 34–48. https://doi.org/10.1007/978-3-030-39540-7_3

Jain, R., Aagja, J., & Bagdare, S. (2017). Customer experience – a review and research agenda. In *Journal of Service Theory and Practice* (Vol. 27, Issue 3, pp. 642–662). Emerald Group Publishing Ltd. https://doi.org/10.1108/JSTP-03-2015-0064

Johnson, M. D., Anderson, E. W., & Fornell, C. (1995). Rational and Adaptive Performance

Expectations in a Customer Satisfaction Framework. *Journal of Consumer Research*, *21*(4), 695. https://doi.org/10.1086/209428

Johnston, R. (1995). The Determinants of Service Quality: Satisfiers and Dissatisfiers. In *International Journal of Service Industry Management* (Vol. 6, Issue 5, pp. 53–71). MCB UP Ltd. https://doi.org/10.1108/09564239510101536

Jowett, S., Yang, X., & Lorimer, R. (2012). The Role of Personality, Empathy, and Satisfaction with Instruction within the Context of the Coach-Athlete Relationship. *Undefined*.

Keeling, K., McGoldrick, P., & Beatty, S. (2010). Avatars as salespeople: Communication style, trust, and intentions. *Journal of Business Research*, *63*(8), 793–800. https://doi.org/10.1016/j.jbusres.2008.12.015

Kerem, E., Fishman, N., & Josselson, R. (2001). The Experience of Empathy in Everyday Relationships: Cognitive and Affective Elements. *Journal of Social and Personal Relationships*, *18*(5), 709–729. https://doi.org/10.1177/0265407501185008

Khedhaouria, A., & Beldi, A. (2014). Perceived enjoyment and the effect of gender on continuance intention for mobile internet services. *International Journal of Technology and Human Interaction*, *10*(2), 1–20. https://doi.org/10.4018/ijthi.2014040101

Kim, S. S., Kaplowitz, S., & Johnston, M. V. (2004). The effects of physician empathy on patient satisfaction and compliance. *Evaluation and the Health Professions*, *27*(3), 237–251. https://doi.org/10.1177/0163278704267037

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information and Management*, *43*(6), 740–755. https://doi.org/10.1016/j.im.2006.05.003

Kirmani, A., Hamilton, R. W., Thompson, D. V., & Lantzy, S. (2017). Doing well versus doing good: The differential effect of underdog positioning on moral and competent service providers. *Journal of Marketing*, *81*(1), 103–117. https://doi.org/10.1509/jm.15.0369

Klein, J., Moon, Y., & Picard, R. W. (2002). This Computer Responds to User Frustration Theory, Design, Results, and Implications. *Undefined*.

Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. *DIS 2017 - Proceedings of the 2017 ACM Conference on Designing Interactive Systems*, 555–565. https://doi.org/10.1145/3064663.3064672

KNAW, NFU, NWO, TO2-federatie, Vereniging Hogescholen, & VSNU. (2018). *Netherlands Code of Conduct for Research Integrity*. https://doi.org/10.17026/dans-2cj-nvwu

Kocaballi, A. B., Quiroz, J. C., Rezazadegan, D., Berkovsky, S., Magrabi, F., Coiera, E., & Laranjo, L. (2020). Responses of conversational agents to health and lifestyle prompts: Investigation of appropriateness and presentation structures. *Journal of Medical Internet Research*, *22*(2). https://doi.org/10.2196/15823

Konijn, E. A., & Hoorn, J. F. (2020). Differential facial articulacy in robots and humans elicit different levels of responsiveness, empathy, and projected feelings. *Robotics*, *9*(4), 1–17. https://doi.org/10.3390/robotics9040092

Kranzbühler, A.-M., Kleijnen, M. H. P., Morgan, R. E., & Teerling, M. (2018). The Multilevel Nature of Customer Experience Research: An Integrative Review and Research Agenda. *International Journal of Management Reviews*, *20*(2), 433–456. https://doi.org/10.1111/ijmr.12140

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, *80*(6), 69–96. https://doi.org/10.1509/jm.15.0420

Liebrecht, C., Sander, L., & van Hooijdonk, C. (2021). Too Informal? How a Chatbot's Communication Style Affects Brand Attitude and Quality of Interaction. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12604 LNCS*, 16–31. https://doi.org/10.1007/978-3-030-68288-0_2

Liu, B., & Sundar, S. S. (2018). Should Machines Express Sympathy and Empathy? Experiments with a Health Advice Chatbot. *Cyberpsychology, Behavior, and Social Networking*, *21*(10), 625–636. https://doi.org/10.1089/cyber.2018.0110

McLean, G., & Osei-Frimpong, K. (2019). Chat now… Examining the variables influencing the use of online live chat. *Technological Forecasting and Social Change*, *146*, 55–67. https://doi.org/10.1016/j.techfore.2019.05.017

Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality*, *40*(4), 525–543. https://doi.org/10.1111/j.1467-6494.1972.tb00078.x

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6–27. https://doi.org/10.1037/met0000086

Moon, M. A., Khalid, M. J., Awan, H. M., Attiq, S., Rasool, H., & Kiran, M. (2017). Percepciones de los consumidores sobre los atributos funcionales y hedonistas de las páginas web, e intenciones de compra online: visión de la actitud cognitivo-afectiva. *Spanish Journal of Marketing - ESIC*, *21*(2), 73–88. https://doi.org/10.1016/j.sjme.2017.07.001

Mordor Intelligence. (2021). *Chatbot Market - Growth, Trends, COVID-19 Impact, and Forecasts (2021 - 2026)*. https://www.researchandmarkets.com/reports/4622740/chatbot-market-growth-trends-covid-19-impact

Morgan, N. A., & Rego, L. L. (2006). The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance. *Marketing Science*, *25*(5), 426. https://doi.org/10.1287/mksc.1050.0180

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nejad, L. M., Wertheim, E. H., & Greenwood, K. M. (2004). Predicting Dieting Behavior by Using, Modifying, and Extending the Theory of Planned Behavior. *Journal of Applied Social Psychology*, *34*(10), 2099–2131. https://doi.org/10.1111/j.1559-1816.2004.tb02692.x

Nguyen, T. N. Q., Tran, Q. H. M., & Chylinski, M. (2020). Empathy and delight in a personal service setting. *Australasian Marketing Journal*, *28*(1), 11–17. https://doi.org/10.1016/j.ausmj.2019.08.003

Nomura, T. (2019). Empathy as Signalling Feedback Between Humanoid Robots and Humans. In *Humanoid Robotics: A Reference* (pp. 2337–2346). Springer Netherlands. https://doi.org/10.1007/978-94-007-6046-2_133

Oliver, R. L., Rust, R. T., & Varki, S. (1997). Customer delight: Foundations, findings, and managerial insight. *Journal of Retailing*, *73*(3), 311–336. https://doi.org/10.1016/S0022-4359(97)90021-X

Overby, J. W., & Lee, E. J. (2006). The effects of utilitarian and hedonic online shopping value on consumer preference and intentions. *Journal of Business Research*, *59*(10–11), 1160–1166. https://doi.org/10.1016/j.jbusres.2006.03.008

Paiva, A., Leite, I., Boukricha, H., & Wachsmuth, I. (2017). Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems*, *7*(3). https://doi.org/10.1145/2912150

Palmer, A. (2010). Customer experience management: A critical review of an emerging idea. *Journal of Services Marketing*, *24*(3), 196–208. https://doi.org/10.1108/08876041011040604

Parasuraman, A., Berry, L. L., & Zeuthaml, V. A. (1991). Retailing : critical concepts. 3,2. Retail practices and operations - Google Books. In *Journal of Retailing* (Vol. 4).

Park, J., & Stoel, L. (2005). Effect of brand familiarity, experience and information on online apparel purchase. *International Journal of Retail and Distribution Management*, *33*(2), 148–160. https://doi.org/10.1108/09590550510581476

Partala, T., & Surakka, V. (2004). The effects of affective interventions in human-computer interaction. *Interacting with Computers*, *16*(2), 295–309. https://doi.org/10.1016/j.intcom.2003.12.001

Peterson, R. A., & Wilson, W. R. (1992). Measuring customer satisfaction: Fact and artifact. *Journal of the Academy of Marketing Science*, *20*(1), 61–71. https://doi.org/10.1007/BF02723476

Prendinger, H., Becker, C., & Ishizuka, M. (2006). A study in users' physiological response to an empathic interface agent. *International Journal of Humanoid Robotics*, *3*(3), 371–391. https://doi.org/10.1142/S0219843606000801

Qualtrics online survey tool. (n.d.). Retrieved May 30, 2021, from https://www.qualtrics.com/free-account/

Reichheld, F. F. (2003). *The One Number You Need to Grow*. www.hbr.org

Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of Personality Assessment*, *93*(1), 84–95. https://doi.org/10.1080/00223891.2010.528484

Rese, A., Ganster, L., & Baier, D. (2020). Chatbots in retailers' customer communication: How to measure their acceptance? *Journal of Retailing and Consumer Services*, *56*, 102176. https://doi.org/10.1016/j.jretconser.2020.102176

Rietz, T., Benke, I., & Maedche, A. (2019). The Impact of Anthropomorphic and Functional Chatbot Design Features in Enterprise Collaboration Systems on User Acceptance. *Wirtschaftsinformatik 2019 Proceedings*. https://aisel.aisnet.org/wi2019/track13/papers/7

Rodrigues, S. H., Mascarenhas, S. F., Dias, J., & Paiva, A. (2009). "I can feel it too!":

Emergent empathic reactions between synthetic characters. *Proceedings - 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*. https://doi.org/10.1109/ACII.2009.5349570

Rosenthal-Von Der Pütten, A. M., Schulte, F. P., Eimler, S. C., Sobieraj, S., Hoffmann, L., Maderwald, S., Brand, M., & Krämer, N. C. (2014). Investigations on empathy towards humans and robots using fMRI. *Computers in Human Behavior*, *33*, 201–212. https://doi.org/10.1016/j.chb.2014.01.004

Rossi, S., Conti, D., Garramone, F., Santangelo, G., Staffa, M., Varrasi, S., & Di Nuovo, A. (2020). The role of personality factors and empathy in the acceptance and performance of a social robot for psychometric evaluations. *Robotics*, *9*(2). https://doi.org/10.3390/ROBOTICS9020039

Ryu, K., Han, H., & Jang, S. S. (2010). Relationships among hedonic and utilitarian values, satisfaction and behavioral intentions in the fast-casual restaurant industry. *International Journal of Contemporary Hospitality Management*, *22*(3), 416–432. https://doi.org/10.1108/09596111011035981

Rzepka, C., Berger, B., & Hess, T. (2020). Why Another Customer Channel? Consumers' Perceived Benefits and Costs of Voice Commerce. *Proceedings of the 53rd Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2020.499

Salminen, J., Jung, S. gyo, Kamel, A. M. S., Santos, J. M., & Jansen, B. J. (2020). Using artificially generated pictures in customer-facing systems: an evaluation study with data-driven personas. *Behaviour and Information Technology*. https://doi.org/10.1080/0144929X.2020.1838610

Schmetkamp, S. (2020). Understanding A.I. — Can and Should we Empathize with Robots? *Review of Philosophy and Psychology*, *11*(4), 881–897. https://doi.org/10.1007/s13164-020-00473-x

Schmitt, B. (1999). Experiential Marketing. *Journal of Marketing Management*, *15*(1–3), 53–67. https://doi.org/10.1362/026725799784870496

Shaw, C., & Ivens, J. (2002). Building Great Customer Experiences. In *Building Great Customer Experiences*. Palgrave Macmillan UK. https://doi.org/10.1057/9780230554719

Shrigley, R. L., Koballa, T. R., & Simpson, R. D. (1988). Defining attitude for science educators. *Journal of Research in Science Teaching*, *25*(8), 659–678. https://doi.org/10.1002/tea.3660250805

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445. https://doi.org/10.1037/1082-989X.7.4.422

Sidaoui, K., Jaakkola, M., & Burton, J. (2020). AI feel you: customer experience assessment via chatbot interviews. *Journal of Service Management*, *31*(4), 745–766. https://doi.org/10.1108/JOSM-11-2019-0341

Sierra Rativa, A., Postma, M., & Van Zaanen, M. (2020). The Influence of Game Character Appearance on Empathy and Immersion: Virtual Non-Robotic Versus Robotic Animals. *Simulation and Gaming*, *51*(5), 685–711. https://doi.org/10.1177/1046878120926694

Spring, T., Casas, J., Daher, K., Mugellini, E., & Khaled, O. A. (2019). Empathic Response Generation in Chatbots. In *Proceedings of 4th Swiss Text Analytics Conference*

(SwissText 2019), 18-19 June 2019, Wintherthur, Switzerland. 18-19 June 2019. https://comprop.oii.ox.ac.uk

Stein, J. P., & Ohler, P. (2017). Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of human-like characters in a virtual reality setting. *Cognition*, *160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010

Strong, K. C., Ringer, R. C., & Taylor, S. A. (2001). The rules of stakeholder satisfaction. *Journal of Business Ethics*, *32*(3), 219–230. https://doi.org/10.1023/A:1010714703936

Suryandari, R. T., & Paswan, A. K. (2014). Online customer service and retail type-product congruence. *Journal of Retailing and Consumer Services*, *21*(1), 69–76. https://doi.org/10.1016/j.jretconser.2013.08.004

Swanson, G. E. (1978). Travels Trough Inner Space: Family Structure and Openness to Absorbing Experiences. *American Journal of Sociology*, *83*(4), 890–919. https://doi.org/10.1086/226636

Swiderska, A., & Küster, D. (2018). Avatars in Pain: Visible Harm Enhances Mind Perception in Humans and Robots. *Perception*, *47*(12), 1139–1152. https://doi.org/10.1177/0301006618809919

Tan, A. H. T., Muskat, B., & Johns, R. (2019). The role of empathy in the service experience. *Journal of Service Theory and Practice*, *29*(2), 142–164. https://doi.org/10.1108/JSTP-10-2018-0221

Turley, D., & O'Donohoe, S. (2017). Mortality, morality and the marketplace: empathetic improvisation and the double duty of care in service encounters with bereaved consumers. *Consumption Markets and Culture*, *20*(5), 456–476. https://doi.org/10.1080/10253866.2017.1367679

Vennix Jac. (2019). *Research methodology An introduction to scientific thinking and practice* (Vol. 1). Pearson Benelux B.V.

Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual Customer Service Agents: Using Social Presence and Personalization to Shape Online Service Encounters. *Journal of Computer-Mediated Communication*, *19*(3), 529–545. https://doi.org/10.1111/jcc4.12066

Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). Customer Experience Creation: Determinants, Dynamics and Management Strategies. *Journal of Retailing*, *85*(1), 31–41. https://doi.org/10.1016/j.jretai.2008.11.001

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and utilitarian dimensions of consumer attitude. In *Journal of Marketing Research* (Vol. 40, Issue 3, pp. 310–320). https://doi.org/10.1509/jmkr.40.3.310.19238

Wang, L. C., Baker, J., Wagner, J. A., & Wakefield, K. (2007). Can A Retail Web Site be Social? *Journal of Marketing*, *71*(3), 143–157. https://doi.org/10.1509/jmkg.71.3.143

Wiese, E., Metta, G., & Wykowska, A. (2017). Robots as intentional agents: Using neuroscientific methods to make robots appear more social. In *Frontiers in Psychology* (Vol. 8, Issue OCT, p. 1663). Frontiers Media S.A. https://doi.org/10.3389/fpsyg.2017.01663

Wieseke, J., Geigenmüller, A., & Kraus, F. (2012). On the Role of Empathy in Customer-Employee Interactions. *Journal of Service Research*, *15*(3), 316–331. https://doi.org/10.1177/1094670512439743

Wilde, P., & Evans, A. (2019). Empathy at play: Embodying posthuman subjectivities in gaming. *Convergence*, *25*(5–6), 791–806. https://doi.org/10.1177/1354856517709987

Wykowska, A., Chaminade, T., & Cheng, G. (2016). Embodied artificial agents for understanding human social cognition. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 371, Issue 1693). Royal Society of London. https://doi.org/10.1098/rstb.2015.0375

Xu, K., & Lombard, M. (2017). Persuasive computing: Feeling peer pressure from multiple computer agents. *Computers in Human Behavior*, *74*, 152–162. https://doi.org/10.1016/j.chb.2017.04.043

Yi, Y. (1989). *A Critical review of consumer satisfaction*. http://deepblue.lib.umich.edu/handle/2027.42/36290

Zamora, J. (2017). I'm Sorry, Dave, i'm afraid i can't do that: Chatbot perception and expectations. *HAI 2017 - Proceedings of the 5th International Conference on Human Agent Interaction*, 253–260. https://doi.org/10.1145/3125739.3125766

Zhou, H., Huang, M., Zhang, T., Zhu, X., & Liu, B. (2017). Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, 730–738. http://arxiv.org/abs/1704.01074

Zhou, L., Yang, Z., & Hui, M. K. (2010). Non-local or local brands? A multi-level investigation into confidence in brand origin identification and its strategic implications. *Journal of the Academy of Marketing Science*, *38*(2), 202–218. https://doi.org/10.1007/s11747-009-0153-1

Zhou, Z., Fang, Y., Vogel, D., Jin, X. L., & Zhang, X. (2012). Attracted to or locked in? predicting continuance intention in social virtual world services. *Journal of Management Information Systems*, *29*(1), 273–306. https://doi.org/10.2753/MIS0742-1222290108

Złotowski, J., Sumioka, H., Nishio, S., Glas, D. F., Bartneck, C., & Ishiguro, H. (2016). Appearance of a robot affects the impact of its behaviour on perceived trustworthiness and empathy. *Paladyn*, *7*(1), 55–66. https://doi.org/10.1515/pjbr-2016-0005

Appendix

Appendix I

Robot Perceived Empathy (RoPE) scale

Perceived Affective Empathy items	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
	1	2	3	4	5
No matter what I tell about myself, the chatbot acts just the same. (Reversed)					
The chatbot encourages me.					
The chatbot praises me when I have done something well.					
The chatbot helps me when I need it.					
The chatbot comforts me when I am upset.					
The chatbot's response to me is so fixed and automatic that I do not get through to it. (Reversed)					

Note: subtracted and adapted from Charrier et al.(2019).

Appendix II

HED/UT scale of attitudes towards mobile information services

	1	2	3	4	5	
Dull						Exciting
Boring						Fun
Serious						Playful
Unamusing						Amusing
Cheerless						Cheerful
Useless						Usefull
Impractical						Practical
Unneccessary						Neccessary
Unfunctional						Functional
Unhelpful						Helpful
Inefficient						Inefficient

Note : The scale is developed by Heijden & Sørensen (2003)