

User Perception of Chatbots in a Mobile Commerce Environment

NORBERTO ALMEIDA DE ANDRADE

norbertofatecsp@hotmail.com

ABSTRACT

In the contemporary landscape, artificial intelligence (AI) is rapidly growing, and organizations are continuously adopting smart technologies. Coupled with chatbots, AI is seen as a critical element in digital transformation and has the power to completely reshape and transform the way customers communicate with organizations. However, there are few empirical investigations on the impact chatbots have on their users. In this context, this article investigates the implementation of interaction heuristics, language, and anthropomorphism in chatbot applications, using a 2x3 experimental research design with mobile commerce videos. The study explores to what extent heuristics related to chatbots can influence perceptions of trust, satisfaction, and purchase intention. The findings of this research confirmed that the use of human-like images and language increases user satisfaction. Social presence in the online environment through anthropomorphic features can also create greater satisfaction. Therefore, organizations, developers, designers, and marketers should focus on choosing a human-like appearance as well as anthropomorphic language cues to enhance satisfaction and provide a better overall user experience.

Keywords: Chatbots; Artificial Intelligence; Mobile Commerce; Consumer Behavior

1. INTRODUCTION

Mobile commerce, also known as m-commerce, refers to the use of wireless portable devices such as cell phones and tablets to conduct online commercial transactions, including buying and selling products, online banking, and bill payments (Tong, Luo & Xu, 2019).

The activity of m-commerce is on the rise, and according to the Mobile Marketing Association (MMA), projections indicated a tenfold increase in investments across all mobile marketing categories between 2019 and 2022. Consumers are purchasing more products and services through mobile devices, reaching a revenue of \$1.7 trillion in 2023. The proportion of mobile devices has consistently increased over time, from 43% in 2018 to a projected 63% by 2028 (Statista, 2024).

The evolution of conversational user interfaces, digital customer service channels, and generative AI will significantly transform customer service and support by 2028 (Gartner, 2023). Retailers are developing and intensifying the use of virtual assistants, such as chatbots and voice bots, with the aim of reaching and informing a larger number of customers about promotions, product launches, and seasonal offers. The level and quality of interaction through chatbots establish a closer and more humanized connection with consumers (Capgemini, 2019).

AI was long regarded as a futuristic imagination or theoretical construct (Kumar *et al.*, 2016). However, the development of AI is accelerating rapidly and is transforming nearly everything connected to human life. In fact, robots or autonomous systems are progressively emerging and have already begun to increasingly replace human labor (Jonke & Volkwein, 2018). This future scenario of intelligent machines and bots that function and react like humans previously existed only as a theoretical possibility. Currently, the development of chatbots is more than present in many sectors and is seen as a popular trend (Davenport *et al.*, 2019).

Technological development has driven and facilitated productive discussions about the use of AI. Despite being in the early stages of implementing models like ChatGPT and Google Gemini, it is possible to perceive the power that AI solutions can generate as they become ubiquitous (Krasodonski *et al.*, 2024)

The AI evolution has led various innovators to want to incorporate emerging technology opportunities in their respective fields. One of these fields where the system has been used is in the development of chatbot applications. Chatbots are described as simulating online interactions with humans, while consulting a database created by natural language input (Kumar *et al.*, 2016).

Sterne (2017) defines chatbots as an application of an AI program that mimics conversations and interactions

with users. The fields of application for chatbots are numerous. However, one of the most popular fields is chatbots for purchasing tickets and searching for or buying products online. Many companies have already recognized this trend and have followed suit by implementing chatbots in their services (Zarouali *et al.*, 2018).

According to the Mobile Marketing Association (MMA) in 2018, more than 54,000 chatbots were already available and active on Facebook's messaging app (Statista, 2017). Major Brazilian brands have begun offering chatbots as a customer service, such as Magazine Luiza, Via Varejo, and Bradesco, to be available 24/7 for their customers (Tong, Luo & Xu, 2019). However, the chatbot trend is not limited to restricted markets, as it is expected to grow on a global scale (Van den Broeck, Zarouali, & Poels, 2019).

One of the biggest advantages of chatbots for companies is their availability, as they can be used 24/7. The use of chatbots as customer service on applications like Facebook is highly profitable and attractive for companies. Most companies can interact with their customers through individual communication on a personal device thanks to chatbots (Jonke & Volkwein, 2018).

AI agents allow companies to create new direct touchpoints with customers and offer communication automation. Automated interaction with customers is not only used to reduce costs but also to increase customer satisfaction (Kumar *et al.*, 2016). Customers can communicate 24/7, regardless of working hours or business hours. Thus, companies can save on personnel costs while still offering customer service. The use of chatbots offers consumers the opportunity to get customer support, obtain personalized recommendations, and click to buy on messaging apps (Zarouali *et al.*, 2018).

Therefore, exceptional benefits for companies include the opportunity to offer customer service 24/7, reduce costs, create direct contact with the customer, and increase customer satisfaction and purchase intention (Shankar, 2016). As the above section describes, it is highly desirable for companies to achieve customer satisfaction and purchase intention. Chatbots can have enormous advantages for companies, however, they must be implemented correctly. Oke *et al.*, (2016) states that users generally distrust intelligent technologies such as chatbots. In other areas of technology, it is well-researched that trust is a critical factor in user adoption of interactive systems (Murphy & Dweck, 2016).

Previous studies have shown that although chatbots were able to provide information and perform basic tasks, they significantly failed to respond empathetically, undermining customer support, impacting user

interaction and satisfaction (Corti & Gillespie, 2019). Researchers Jain *et al.* (2018) noted that when novice users do not understand the capabilities and limitations of chatbots, they have high expectations and consequently become more frustrated when chatbots fail and do not meet expectations.

Other research shows that user trust in chatbots for customer service was affected by factors related to the specific appearance of the chatbot, specifically the quality of its human appearance (Luo *et al.*, 2019). Another significant challenge for chatbot adoption is that interactions with them often do not feel natural or human (Zarouali *et al.*, 2018).

Thus, the primary desire of a user is to experience a natural conversation with a chatbot that seems human (Jonke & Volkwein, 2018). It is suggested to create a chatbot as human-like as possible. Social or human cues, such as language style, can influence and enhance the perception of anthropomorphism (Hofacker, Malthouse & Sultan, 2016).

Similarly, Davenport *et al.*, (2019) states that the quality of a virtual assistant's dialogue can lead to improved customer satisfaction. Solomon *et al.* (2014) argues that not only the language style, referred to as textual information, evokes online reliability, but visual cues also play a crucial role. Visual cues, such as the chatbot's appearance, are design features to make chatbot interactions seem more natural and human-like (Zarouali *et al.*, 2018). Tong, Luo & Xu (2019) suggest that consumers even prefer the female gender in chatbot appearances, although most brands and companies do not use human images or animations and instead create logos for their online services.

In fact, most e-service websites generally lack human appearances, which can harm the purchase intention of potential customers as well as trust development, as online interactions with social presence are believed to be crucial in building customer trust (Solomon *et al.*, 2014).

Therefore, this research will examine the effects of chatbot appearance, i.e., a human, an animated person, and an organizational logo of Magazine Luiza. The Magazine Luiza chatbot, also known as Lu, was created by the retail chain to help with customer service and improve post-sales process agility (Hofacker, Malthouse & Sultan, 2016).

As the primary factor determining chatbot adoption and use is perceived anthropomorphism, two crucial factors will be examined. Alongside chatbot appearance, the perceived humanity of chatbot language will be explored. Thus, this research further examines chatbot language, which can be robotic or human (Sterne, 2017).

The current development of chatbots for mobile messengers is still in its infancy and, due to its novelty, there is currently little or no research on the anthropomorphic features of visual and linguistic chatbots in the mobile messenger interface environment (Tong, Luo & Xu, 2019).

Many companies only use the company logo in chatbot design to interact with their customers. However, theories such as social presence theory suggest that the mere presence of a human image can have a better impact on user engagement with the chatbot. Thus, implementing anthropomorphic cues in terms of chatbot appearance and language may have a better impact on user perception compared to just a logo. Therefore, it is important to know if users trust different types of chatbots depending on the degree of anthropomorphic characteristics of the appearance or language (Davenport *et al.*, 2019). Implementing the correct chatbot design in companies can not only lead to user trust but also increase satisfaction and purchase intention, which are crucial for business success (Zarouali *et al.*, 2018). Consequently, the objective of this research is formulated in the following research questions:

RQ1: To what extent does chatbot appearance influence trust, satisfaction, and purchase intention?

RQ2: To what extent does robotic/human language influence trust, satisfaction, and purchase intention?

RQ3: To what extent do the effects of chatbot appearance on trust, customer satisfaction, and purchase intention depend on the robotic/human language used for interaction?

RQ4: To what extent are the effects of chatbot appearance and robotic/human language on trust, satisfaction, and purchase intention mediated by trust?

This research is divided into several sections. Firstly, chapter two describes a theoretical framework with the dependent variables (trust, satisfaction, purchase intention) and independent variables (chatbot appearance, language) of this research.

The hypotheses are derived from the framework and followed by the research model of this study. Secondly, the research methods and designs are elaborated on in chapter three.

The results of this research are presented in chapter four, followed by chapter five, which presents a discussion of the results. Finally, the limitations of this research and a conclusion are presented.

2. THEORETICAL FRAMEWORK

Anthropomorphism is the attribution of a set of characteristics—both physical and non-physical, behaviors, emotions, and attributes—that are like humans to a

non-human agent or animated object (Epley *et al.*, 2007). Historically, anthropomorphism has significant roots and documentation, with the earliest drawings from 30,000 years ago depicting animals with human-like bodies (Dalton, 2004).

In human-computer interaction, anthropomorphism is typically triggered by anthropomorphic cues, whereby characteristics, behaviors, or design elements that ascribe human qualities to inanimate objects or artificial agents within information technology (IT) are employed. Anthropomorphism could influence IT users (Burgoon *et al.*, 2000; Epley *et al.*, 2007).

Anthropomorphic characteristics developed from anthropomorphic design are multifaceted and include, for example, voice recognition, voice synthesis, and computer-generated graphical representations of human-like faces or bodies, including facial expressions and gestures, among others (Woods, 2015; Li *et al.*, 2016).

However, anthropomorphism is not just about addressing the visual or auditory aspects of interaction in a more human manner but also about making the overall interaction between the user and the computer more human-like. The pursuit of natural language in chatbots' interactions with users is ongoing, and the demand for more contextualized, sensitive, and semantically correct behaviors and responses would promote a more trustworthy experience in interactions and elevate the perception of human similarity (Abdul-Kader & Woods, 2015; Li *et al.*, 2016).

The ramifications of this evolution can be observed in features like exclusive voice recognition, friendly appearance, and personalized user interfaces that we utilize (Haas *et al.*, 2020). The most current technological products are agents or virtual assistants, such as Apple's Siri and Amazon's Alexa, and OpenAI's ChatGPT. Robust e-commerce platforms like Amazon and Alibaba have incorporated chatbot and voice shopping features into their devices, demonstrating the enormous potential of this interaction (Klaus & Zaichkowsky, 2020).

In the United States, only 15% of smart speaker users regularly make purchases via voice commands, indicating a wide array of opportunities (Voicebot.ai, 2021). Although these agents lack a physical body, humans tend to seek interpersonal relationships and social connections through increasingly deeper communications and interactions (Fan *et al.*, 2017) with both animated and inanimate agents.

Only a few empirical studies have been conducted on the human-like appearance of a chatbot and its

anthropomorphic effect (Van den Broeck, Zarouali, & Poels, 2019).

For this reason, this research explores the aspects of anthropomorphic visual features more closely, using a human image, an animated human, and a logo in a mobile-commerce environment. Based on the results of previous sections, the following hypothesis was formulated:

H1: The perception of (a) trust, (b) satisfaction, and (c) purchase intention is higher when people are confronted with a chatbot using a human image compared to people using a chatbot with an animated image or organizational logo.

In the last decade, new human-computer interfaces have emerged, combining numerous human language technologies that allow humans to interact and communicate with computers using spoken or written dialogue to access, create, and process information (Davenport *et al.*, 2019).

These platforms that mimic a conversation with a real human being are called conversational interfaces (CIs). CIs offer users the opportunity to communicate with a computer (or chatbot) in their natural language or, in other words, in human language, rather than a specific syntax command (Jonke & Volkwein, 2018).

Natural language is not only a dependent variable in this research. It is assumed that the chatbot's language affects the direction or strength of the relationship between the chatbot's appearance and the independent variables of trust, satisfaction, and purchase intention (Luo *et al.*, 2019).

In this research, high and low natural language is displayed based on the main characteristics of Van den Broeck, Zarouali, & Poels (2019) to clearly distinguish human and chatbot conversations. These two conditions are named "robotic" for low natural language and "human" for high natural language performance (Zarouali *et al.*, 2018). Based on the results of the previous sections, the following hypotheses were formulated:

H2: The perception of (a) trust, (b) satisfaction, and (c) purchase intention is higher when people are confronted with a chatbot using human language compared to people using a chatbot with robotic language.

H3: The perception of (a) trust, (b) satisfaction, and (c) purchase intention is higher when people are confronted with a chatbot using both human image and language compared to people using a chatbot with a logo or animated image and human language.

H4: The perception of (a) trust, (b) satisfaction, and (c) purchase intention is higher when people are confronted with a chatbot using robotic language and an animated image or logo compared to people using a chatbot with robotic language and human language.

Analyzing the literature on the variables of this article—namely, chatbot appearance and natural language as independent variables, purchase intention, and satisfaction as dependent variables—studies indicate that trust plays a vital role in relation to both dependent and independent variables. Therefore, trust is seen as the mediating variable of this research. Based on the results of the previous section, the following hypotheses were formulated:

H5: The effects of chatbot appearance on (b) satisfaction and (c) purchase intention are mediated by trust.

H6: The effects of natural language on (b) satisfaction and (c) purchase intention are mediated by trust.

H7: Language moderates the impact of chatbot appearance on (a) trust, (b) satisfaction, and (c) purchase intention.

2.1. Research Model

Based on the reviewed literature and previous studies, a research model was developed, illustrated in Figure 1. The objective is to explore the effects of chatbots' appearance and language on trust, purchase intention, and satisfaction.

3. METHODOLOGY

To investigate the effect of chatbot appearance and language, this research employed a 2x3 design (Wierenga & Van der Lans, 2017). The three different conditions vary in the chatbot's appearance: human image, and

animated image, and organizational logo. These are combined with two language factors: robotic or natural (King, 2019). This three-by-two experimental design is shown in Table 1.

The stimuli consisted of six different conversations on the Facebook chatbot interface. Six different conversational interfaces were created for this experiment. The conversational interface displayed a Facebook Messenger interface with three different chatbot appearances and two different types of language (Jonke & Volkwein, 2018).

The first type of chatbot was created with (1) a human appearance, the second with (2) an animated human image, and the third simply used the Magazine Luiza logo (3). Each of the three chatbots displayed a conversational interface in either natural language (4) or robotic language (5). These six different conditions were shown as a video approximately 30 seconds long. One of the six conditions was randomly assigned to the participants.

Before interacting with the chatbot, participants needed to read a scenario. This scenario described the option of using the Lu (Magazine Luiza) chatbot on Facebook to find an appropriate gift for a friend. Participants had to watch one of the six videos in which the conversation with one of the chatbots was displayed. As participants were unable to interact directly with the chatbots themselves, they were asked to imagine being the person interacting with the chatbot in the provided video.

Three different photos were shown in the chatbot conversation. As Davenport *et al.* (2019) found, users prefer female appearances in chatbots, despite many online brand animations or logos. Therefore, the different chatbot appearances were represented as (1) a human, the second image (2) an animated human, and the third image simply used the Magazine Luiza

Figure 1 – 2x3 Model

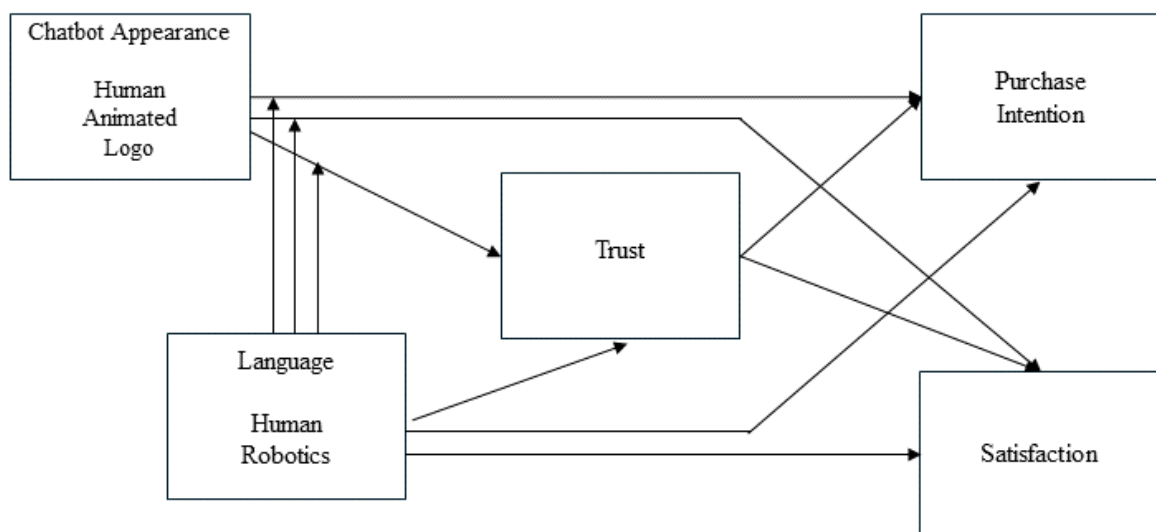


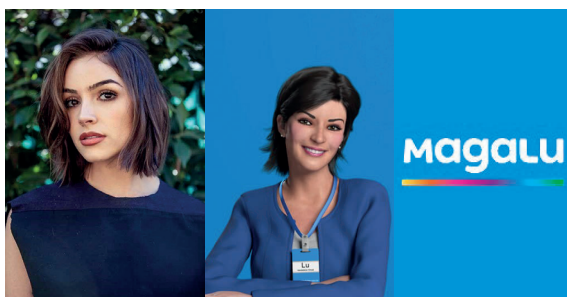
Table 1

2x3 Experimental Design with 6 Conditions

| | | Chatbot Appearance | | |
|----------|------------------|--------------------|--------------|--------------|
| | | Human | Animated | Logo |
| Language | Natural Language | Conversation | Conversation | Conversation |
| | | 1 | 2 | 3 |
| | Robotic Language | Conversation | Conversation | Conversation |
| | | 1 | 2 | 3 |

logo (3). The images used for the chatbot are illustrated in Figure 2.

Figure 2. Images of Chatbot Appearance



1) Human Image 2) Humanized Chatbot 3) Logo

Each of the three chatbots displayed a conversational interface in either natural language (4) or robotic language (5). Van den Broeck, Zarouali, & Poels (2019) identified differences between human conversations and those with chatbots in terms of words per message, words per conversation, vocabulary richness, closed questions, and word exclusivity. Additionally, people interacted with chatbots for longer periods

but with shorter messages compared to interactions with another human.

These elements were implemented in the two different types of conversation. An example of these two conditions is presented in Figure 3, showcasing the humanized appearance of the chatbot.

After participants viewed one of the six conditions, they completed an online questionnaire to measure the effects. For this study, a manipulation check was conducted as an indicator of the internal validity of the experiment (Hofacker, Malthouse & Sultan, 2016).

The manipulation check was performed to investigate whether the manipulation of chatbot appearance and language was effective. First, a one-way ANOVA and Post Hoc test were conducted for the chatbot appearance, followed by a t-test for the language (Fávero & Belfiore, 2017).

In this experiment, participants responded to 7 items on a 5-point Likert scale regarding the appearance of the chatbot. The semantic scale ranged from 1 'strongly

Figure 3. Human and Robotic Language Conversation with Animated Image.



agree' to 5 'strongly disagree.' Thus, a lower average value indicated a higher perceived anthropomorphism of the chatbot's appearance (Malhotra, 2015).

To determine if there were statistically significant differences between the means of the three chatbot appearance groups—human, animated, and logo—a one-way ANOVA was conducted. Initially, the seven chatbot appearance items were combined into a new variable in SPSS. Subsequently, a one-way ANOVA was performed to check for significant differences between the meanings of the three groups. Additionally, a Post Hoc test was executed to explore where the differences occurred between the groups (Marôco, 2018).

The results from the ANOVA test indicated a significant effect on the three conditions. Specifically, the values showed significant differences $M = 1.91$, with $F(2, 221) = 13.13$, $p < 0.001$ among the three groups. Moreover, a Post Hoc test confirmed where the differences occurred between the groups (Wierenga & Van der Lans, 2017).

The test results indicate a significant difference between the human group and the animated group, $p < 0.001$, as well as between the human group and the logo group, $p < 0.001$. However, the Post Hoc test also revealed no significant difference between the animated and logo groups ($p = 0.954$). The results confirm the assumption that a chatbot with a human image is perceived as more human than the animated image and the Magazine Luiza logo (Fávero & Belfiore, 2017).

In addition to the chatbot appearance groups, this research created two different language groups (human and robotic) for the independent variable 'language.' Participants had to respond to 7 items on a 5-point Likert scale regarding the perceived anthropomorphism of the chatbot's language (Malhotra, 2015).

Similarly, the semantic scale ranged from 1 'strongly agree' to 5 'strongly disagree,' and a lower average value indicated a higher perception of anthropomorphism. An independent samples t-test was conducted to confirm whether there were significant differences between the means of the two language groups (Marôco, 2018).

The t-test results showed significant differences between robotic language ($M = 2.44$, $SD = 0.91$) and human language ($M = 1.84$, $SD = 0.91$), with $t(222) = 5.52$, $p < 0.001$. These statistically significant results suggest that respondents recognized the two different language styles in the study. However, the difference between the two language group means is not as large as expected (Wierenga & Van der Lans, 2017).

3.1. Participants

For this experiment, a total of 267 participants completed the questionnaire. However, 43 questionnaires were excluded due to incomplete responses or participants not meeting the criteria for this research.

Thus, the dataset used in this study consists of 224 respondents.

The average age of the participants was $M = 24.5$ years, $SD = 7.5$. Additionally, most respondents, specifically 74, hold a high school diploma, followed by 61 individuals with a bachelor's degree. Furthermore, 51 respondents have some college credit but no degree, and 19 hold a master's degree. Finally, 12 participants have an associate's degree, 2 individuals have less than a high school diploma, and 2 have a doctoral degree. Moreover, 83% use Facebook Messenger and have interacted with a chatbot before.

The online survey was created using Qualtrics, and the chatbot interaction videos were designed with the online tool Bot preview. The main study survey was distributed through online channels such as email, social media (Facebook, Instagram, and WhatsApp) (Tong, Luo & Xu, 2019).

Subsequently, participants were required to read a scenario where they imagined being in a situation where they wanted to buy products on an e-commerce site for electronics for a friend's wedding. Participants were asked to imagine being the person interacting with the chatbot and to watch the video carefully.

Afterward, one of the six chatbot interaction videos was randomly assigned to the participant. During the video, the chatbot appeared with a human image, animated image, or logo. Along with the three different chatbot appearance options, the chatbot engaged in either robotic or human-like language.

Following the video interaction, participants were asked to complete the questionnaire. The survey included sets of questions regarding the chatbot's appearance, language, satisfaction, purchase intention, and trust (although the trust data were not used for further analysis).

To investigate the performance of the items in this survey concerning other variables, a construct validity test was conducted through factor analysis, explained variance, eigenvalues, and Cronbach's alpha calculation to assess reliability (Fávero & Belfiore, 2017).

To verify the validity of the study, a factor analysis was performed. A total of 24 items were analyzed, divided into 5 factors, which correspond to two dependent variables and three independent variables (Malhotra, 2015). The aim was to determine whether the variables measured what they were intended to measure and whether the five factors were distributed across the expected five constructs.

Table 2 shows the SPSS factor analysis, with the dependent and independent variables represented. The items for the chatbot appearance, language, trust, satisfaction, and purchase intention variables were

portrayed. Items for the chatbot appearance, language, and purchase intention variables ended up in one factor column, indicating that these items measured what they were supposed to measure. However, items for the satisfaction and trust variables were found in the same column, suggesting that the two variables measured the same factor instead of two separate ones (Fávero & Belfiore, 2017).

Trust and satisfaction were strongly correlated within their measures; however, it is not conceptually possible to merge them into a single construction. It was hypothesized that trust would have an expected mediating effect, which must be rejected due to the elimination of the variable. Furthermore, all hypotheses containing the effects of trust must be rejected (Solomon *et al.*, 2014).

Additionally, the explained variance for all variables is 68.33%. Higher percentages of explained variance indicate a stronger association. Thus, the variance explained in this study had a relatively high score and is therefore acceptable. Although the explained variance did not exceed 70%, it can still be significant, as it indicates that the regression model has statistically significant explanatory power (Fávero & Belfiore, 2017).

Eigenvalues indicate the strength of transformation in a specific direction. Each eigenvalue for each factor in this study is greater than 1, indicating that the items in this survey are valid (Marôco, 2018).

The reliability of this research was tested by calculating Cronbach's Alpha for the variables to assess internal consistency, i.e., how closely related a set of items is as a group (Wierenga & Van der Lans, 2017).

A summary of Cronbach's Alpha for each variable is shown in Table 2 of the validity factor analysis. Each Cronbach's Alpha for the four variables—chatbot appearance, language, satisfaction, and purchase intention—is above 0.7, suggesting that the items have relatively high internal consistency and, therefore, confirm an acceptable level (Fávero & Belfiore, 2017).

4. RESULTS

In this study, six different conditions were designed to examine the influence of these conditions on the dependent variables of satisfaction, purchase intention, and confidence (though confidence was eliminated and is no longer used from this point onward) (Murphy & Dweck, 2016).

First, a Multivariate Analysis of Variance (MANOVA) was conducted to analyze the survey data involving more than one dependent variable at a time (Malhotra, 2015). The MANOVA explored the hypotheses about the effects of the two independent variables (chatbot appearance and language) on the two dependent variables (satisfaction and purchase intention). Additionally,

the main effects of visual appearance, language, and the chatbot on satisfaction and purchase intention were elaborated (Marôco, 2018).

The following section provides a more detailed explanation of the non-existent interaction effect. Finally, an overview of the supported or rejected hypotheses is presented. This study used MANOVA to examine the different effects of the independent variables (chatbot appearance and language) on the dependent variables (confidence, satisfaction, and purchase intention). The results are presented separately for each dependent variable (Fávero & Belfiore, 2017).

Before analyzing the MANOVA results, Wilks' Lambda was calculated to check the overall effect between the two independent variables (Malhotra, 2015). Table 3 shows the descriptive statistics of the independent variables, chatbot appearance and language.

Observing the first main effect of this study, chatbot language, it can be stated that there is a main effect of chatbot language, with $\Lambda = 0.952$, $F(3, 216) = 3.625$, $p = 0.014$. Additionally, the second main effect, chatbot appearance, has a significant value of $\Lambda = 0.945$, $F(6, 432) = 2.084$, $p = 0.054$, which means there is no significant main effect of chatbot appearance. However, the p-value is nearly significant ($p = 0.054$), and therefore will be considered in this analysis and not discarded, even though it is an insignificant result (Fávero & Belfiore, 2017).

Regarding the interaction between the two main effects, language * chatbot appearance, it indicates that there is no significant interaction effect between the two groups, with $\Lambda = 0.961$, $F(6, 432) = 1.444$, $p = 0.197$. The hypotheses of this research expected that the language variable would moderate the impact of chatbot appearance (Marôco, 2018). However, Table 3 shows that there is no significant interaction effect between the two independent variables and, therefore, no moderation effect as was expected. For this reason, Hypothesis 7 (H7) must be rejected.

The main effects of the independent variables, chatbot appearance and language, on the dependent variables (satisfaction, purchase intention) were further explored in the following sections. As shown in Table 4, the effects between the independent variables (chatbot appearance and language) and the dependent variables (satisfaction, purchase intention) were measured. The results indicate a main effect between language and satisfaction, with $(p = 0.033)$ ($(F = 4.612)$), and between chatbot appearance and satisfaction, with $(p = 0.016)$ ($(F = 4.192)$) (Wierenga & Van der Lans, 2017).

The satisfaction items in this study were measured using a 5-point Likert scale, ranging from 1 = 'strongly agree' to 5 = 'strongly disagree'. The Bonferroni results from the MANOVA reveal a difference among

Table 2

Factor Analysis of Validity

| Factor 1: Chatbot Appearance ($\alpha = .914$) | 1 | 2 | 3 | 4 |
|--|----------|----------|----------|----------|
| The impression of the chatbot's image seemed lively. | | .81 | | |
| The impression of the chatbot's image seemed animated. | | .78 | | |
| The impression of the chatbot's image seemed natural. | | .79 | | |
| The impression of the chatbot's image seemed interactive. | | .71 | | |
| My impression of the chatbot's image seemed natural. | | .81 | | |
| My impression of the chatbot's image seemed human. | | .79 | | |
| My impression of the chatbot's image seemed realistic. | | .83 | | |
| Factor 2: Chatbot Language ($\alpha = .933$) | 1 | 2 | 3 | 4 |
| My impression of the chatbot's language seemed natural. | .85 | | | |
| My impression of the chatbot's language seemed human. | .85 | | | |
| My impression of the chatbot's language seemed realistic. | .85 | | | |
| My impression of the chatbot's language seemed lively. | .82 | | | |
| My impression of the chatbot's language seemed animated. | .76 | | | |
| My impression of the chatbot's language seemed natural. | .80 | | | |
| My impression of the chatbot's language seemed interactive. | .69 | | | |
| Factor 3: Trust ($\alpha = .840$) | 1 | 2 | 3 | 4 |
| The chatbot was trustworthy. | | | .77 | |
| The chatbot was authentic. | | | .68 | |
| The chatbot was competent. | | | .79 | |
| The chatbot was agile.. | | | .79 | |
| Factor 4: Satisfaction ($\alpha = .809$) | 1 | 2 | 3 | 4 |
| I am satisfied with the chatbot's interface. | | | .57 | |
| My choice to ask the chatbot about a product was sensible. | | | .63 | |
| I am satisfied with the way the chatbot assisted me. | | | .70 | |
| Factor 5: Purchase Intention ($\alpha = .884$) | 1 | 2 | 3 | 4 |
| It is very likely that I will purchase via mobile. | | | | .76 |
| I intend to purchase a product within 3 months. | | | | .91 |
| I intend to purchase a product within 6 months. | | | | .90 |

the three chatbot groups (human, animated, logo). Survey respondents were most satisfied with the human appearance chatbot ($M = 1.84, SD = 0.085$), followed by the animated image ($M = 1.95, SD = 0.081$), and lastly, the logo ($M = 2.95, SD = 0.082$), with $(F(2, 218) = 4.192)$, $(p = 0.016)$ (Fávero & Belfiore, 2017).

It was hypothesized that participants would have a higher perception of satisfaction when confronted with a chatbot displaying a human image compared

to those interacting with a chatbot displaying an animated image or organizational logo. Therefore, Hypothesis (H1b) that satisfaction perception is higher when people interact with a chatbot using a human image compared to a chatbot with an animated image or organizational logo, can be confirmed based on the results (Malhotra, 2015).

Respondents in this study were most satisfied with a chatbot using human language ($M = 1.87, SD = 0.067$)

Table 3
Multivariate Test; Descriptive Statistics of Independent Variables

| Effect | | Value | F | Sig. | Partial Eta Squared |
|-----------------|-----------------|-------|-------|------|---------------------|
| Language Group | lambda de Wilks | .952 | 3.625 | .014 | .048 |
| Chatbot Group | lambda de Wilks | .945 | 2.084 | .054 | .028 |
| Language Group* | lambda de Wilks | .961 | 1443 | .197 | .020 |
| Chatbot Group | | | | | |

compared to robotic language ($M = 2.08, SD = 0.067$), with $(F(1, 218) = 4.612)$, $(p = 0.033)$. It was hypothesized that participants would perceive higher satisfaction when interacting with a chatbot using human language compared to those interacting with a chatbot using robotic language. Therefore, Hypothesis H2b can be confirmed according to the investigated results (Marôco, 2018).

Table 4 shows that the results for purchase intention indicate no main effect between language and purchase intention, with $(p = 0.137)$ $(F = 2.223)$. Similarly, there is no main effect between chatbot appearance and purchase intention, with $(p = 0.226)$ $(F = 1.346)$ (Wierenga & Van der Lans, 2017).

It was hypothesized that participants would have a higher purchase intention when confronted with a chatbot displaying a human image compared to those interacting with a chatbot displaying an animated image or organizational logo (Murphy & Dweck, 2016). It was also expected that participants would have a higher purchase intention when interacting with a chatbot using natural language compared to those interacting with a chatbot using robotic language. However, since there is no main effect of visual appearance and

chatbot language on purchase intention, Hypotheses H1c and H2c should be rejected (Fávero & Belfiore, 2017).

Although it has been previously claimed that there is no interaction effect between the appearance and language of the chatbot, a MANOVA was conducted to measure the effects of the interaction between the chatbot's appearance and language on the dependent variables (satisfaction, purchase intention).

Examining the interaction between the two main effects (chatbot language and appearance) in Table 4 indicates that there is no significant interaction effect between the chatbot's appearance and language, with $\Lambda = 0.961, F(6, 432) = 1.443, p = 0.197$. Furthermore, Table 5 shows that there is no interaction effect of the chatbot's language appearance on the two different dependent variables with $p = 0.175$ ($F = 1.756$) for satisfaction and $p = 0.288$ ($F = 1.260$) for purchase intention (Malhotra, 2015).

It was expected, based on two hypotheses, that certain combinations of the three chatbot appearances and language would lead to different effects on the dependent variables of satisfaction and purchase intention. However, as the effect of the chatbot's language appearance on the dependent variables is

Table 4
Between-Subjects Effects Test

| Source | Dependent Variables | df | F | Sig. |
|----------------|---------------------|----|-------|------|
| Language Group | Satisfaction | 1 | 4.612 | .033 |
| | Purchase Intention | 1 | 2.246 | .137 |
| Chatbot Group | Satisfaction | 2 | 4.192 | .016 |
| | Purchase Intention | 2 | 1346 | .262 |

insignificant, H3 and H4 must be rejected (Fávero & Belfiore, 2017).

4.1. Overview of Hypotheses

Table 6 provides an overview of the 7 hypotheses and indicates whether they are supported or rejected based on the results of this research.

5. DISCUSSION OF RESULTS

The aim of this study was to gain a better understanding of the effects of using anthropomorphic characteristics in the appearance and language of chatbots, implementing visual or verbal human cues. The effects of trust, satisfaction, and purchase intention concerning online interactions with messenger chatbots were explored. Six different conditions were created to test the effect of anthropomorphism on chatbot appearances (human, animation, and logo) and language (human, robotic) on perceived trust, satisfaction, and purchase intention. Language as a moderator and trust as a mediator were examined.

To address these objectives, four research questions were proposed: (1) To what extent does the chatbot's appearance influence trust, satisfaction, and purchase intention? (2) To what extent does robotic/human language influence trust, satisfaction, and purchase intention? (3) To what extent do the effects of a chatbot's appearance on trust, customer satisfaction, and purchase intention depend on the robotic/human language used for interaction? (4) To what extent are the effects of the chatbot's appearance and robotic/human language on trust, satisfaction, and purchase intention mediated by trust?

These research questions were answered through an experimental design and seven main hypotheses. The most significant findings from this research were the confirmation of the two hypotheses related to human image and language, leading to higher levels of satisfaction among users. The reviewed literature indicated that human-like cues, such as language style or visual appearance, can influence and enhance the perception of anthropomorphism (Jonke & Volkwein, 2018) and thus lead to improved customer satisfaction

(Solomon *et al.*, 2014). Observed studies suggest that implementing anthropomorphic visual cues, using a human image, increases the perception of social presence, which positively influences satisfaction (Murphy & Dweck, 2016).

Based on this literature finding, it was expected that participants in this study would be more satisfied with a chatbot displaying a human image rather than an animation or logo. Additionally, it was anticipated that respondents would be more satisfied with the anthropomorphic linguistic features of the human language group compared to the robotic language group (Jonke & Volkwein, 2018).

The results of this study showed that participants were most satisfied with the chatbot's visual and linguistic anthropomorphic features. Thus, these findings further support the idea of implementing social presence in the online environment through anthropomorphic features to create greater user satisfaction. Moreover, this finding has important implications for chatbot development and design, aiming to create a better user experience. Implementing anthropomorphic cues in chatbots creates better user experience and results in happier customers (Oke *et al.*, 2016).

The reviewed literature indicated that online visual features, such as the chatbot's appearance, alongside textual information, are factors that can lead to a higher purchase intention (Murphy & Dweck, 2016). Therefore, it was expected that human images in the chatbot's appearance would lead to a higher purchase intention. Additionally, this research expected that the human language group would score higher on purchase intention compared to the robotic group (Solomon *et al.*, 2014).

However, the results of this research must reject this assumption, as there is no main effect of the appearance or language of the chatbot on purchase intention. Research has shown that user trust in chatbots was affected by factors related to the specific appearance of the chatbot, specifically the quality of its human appearance, which can be incorporated through different types of chatbot appearances and natural language (Van den Broeck, Zarouali, & Poels, 2019).

Table 5
Between-Subjects Effects Test

| Source | Dependent Variables | df | F | Sig. |
|-----------------|---------------------|----|-------|------|
| Language Group* | Satisfaction | 2 | 1.756 | .175 |
| Chatbot Group | Purchase Intention | 2 | 1.260 | .288 |

Table 6
Overview of Hypotheses

| Hypotheses | | Supported |
|------------|--|-----------|
| H1a | Perceived trust is higher when individuals interact with a chatbot using a human image compared to those interacting with a chatbot using an animated image or organizational logo. | No |
| H1b | Perceived satisfaction is higher when individuals interact with a chatbot using a human image compared to those interacting with a chatbot using an animated image or organizational logo. | Yes |
| H1c | Perceived purchase intention is higher when individuals interact with a chatbot using a human image compared to those interacting with a chatbot using an animated image or organizational logo. | No |
| H2a | Perceived trust is higher when individuals interact with a chatbot using natural language compared to those interacting with a chatbot using machine-like language. | No |
| H2b | Perceived satisfaction is higher when individuals interact with a chatbot using natural language compared to those interacting with a chatbot using machine-like language. | Yes |
| H2c | Perceived purchase intention is higher when individuals interact with a chatbot using natural language compared to those interacting with a chatbot using machine-like language. | No |
| H3a | Perceived satisfaction is higher when individuals interact with a chatbot using a human image and natural language compared to those interacting with a chatbot using a logo or animated image with human-like language. | No |
| H3b | Perceived satisfaction is higher when individuals interact with a chatbot using a human image and natural language compared to those interacting with a chatbot using a logo or animated image with machine-like language. | No |
| H3c | Perceived purchase intention is higher when individuals interact with a chatbot using a human image and natural language compared to those interacting with a chatbot using a logo or animated image with human-like language. | No |
| H4a | Perceived trust is higher when individuals interact with a chatbot using robotic language with an animated image or logo compared to those interacting with a chatbot using robotic language and human language. | No |
| H4b | Perceived satisfaction is higher when individuals interact with a chatbot using robotic language with an animated image or logo compared to those interacting with a chatbot using robotic language and human language. | No |
| H5a | The effects of the chatbot's appearance on satisfaction are mediated by trust. | No |
| H5b | The effects of the chatbot's appearance on purchase intention are mediated by trust. | No |
| H6a | The effects of natural language on satisfaction are mediated by trust. | No |
| H6b | The effects of natural language on purchase intentions are mediated by trust. | No |
| H7a | Language moderates the impact of the chatbot's appearance on trust. | No |
| H7b | Language moderates the impact of the chatbot's appearance on satisfaction. | No |
| H7c | Language moderates the impact of the chatbot's appearance on purchase intention. | No |

Consequently, it was expected that anthropomorphic cues would lead to higher perceptions of trust. However, factor analysis showed that trust results were not statistically significant and were correlated with satisfaction (Fávero & Belfiore, 2017). Conceptually, it was not possible to merge them and, therefore, had to be eliminated. Trust data were no longer used for further analyses and evaluations. One reason for the insignificant results might be that this research used only four trust items to gain insight into predicting participants' initial trust levels in chatbots. This scale was overly simplified to measure online trust (Davenport *et al.*, 2019). Other items and factors should have been considered, such as trust propensity, knowledge-based trust, such as familiarity or brand reputation. Chatbots collect confidential data, and thus perceived security and privacy may have influenced trust perceptions toward the chatbot as well. For this reason, the trust variable could not be further explored (Zarouali *et al.*, 2018).

In turn, anthropomorphic visual and linguistic features resulted in higher satisfaction among the users of this study. This finding supports the theory of social presence and shows that this theory applies similarly to the mobile commerce environment with chatbots. Furthermore, it helps chatbot designers and companies develop mobile commerce chatbots that satisfy customers (Luo *et al.*, 2019).

The visual appearance of chatbots and the importance of the displayed language are seen as crucial synergy. Additionally, previous studies have shown that voice tone had a moderating effect on consumer responses on social media, with an interaction between voice tone and human facial characteristics (Jonke & Volkwein, 2018). Therefore, it was expected that the impact of the chatbot's appearance would depend on the chatbot's language. Further research on this topic needs to be conducted before the association between the chatbot's appearance and chatbot language cues is more clearly understood.

6. CONCLUSION

This study was developed to gain a better understanding of the effects of using human or robotic characteristics in the appearance and language of chatbots in an online environment. Additionally, the effects of the chatbot's appearance and language on trust, satisfaction, and purchase in mobile commerce were explored.

The most notable finding was that respondents were more satisfied with a chatbot using human image and language. However, the implementation of these cues did not increase purchase intention according to the results of this research. Organizations and chatbot designers should seek anthropomorphic cues in the

development and implementation of chatbots to create a better experience for users interacting with intelligent agents.

It is imperative to acknowledge the inherent limitations of this study. The experiment should be conducted in a more realistic and enhanced environment in the future, in terms of design and research measures. Specifically, measurement instruments for purchase intention and trust need to be improved. More robust trust scales with additional items could be used to capture the complexity of online trust. Perceptions of security and privacy could also be analyzed for a comprehensive understanding of the influence of trust in chatbots.

Thus, having real interactions between humans and chatbots would test whether the results would be like those of this research, as the uncanny valley theory argues that many anthropomorphic cues can lead to results contrary to this study. The related limitations indicate that while the study has provided valuable insights into the impact of anthropomorphic characteristics on chatbots, some areas still require future research to improve the validity and applicability of the results in real-world contexts.

REFERENCES

- Abdul-Kader, S. A., & Woods, J. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, 6(7), 72-80.
- Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., & Allspach, L. (2000). Interactivity in human-computer interaction: A study of credibility, understanding, and influence. *Computers in Human Behavior*, 16(6), 553-574. DOI: [https://doi.org/10.1016/S0747-5632\(00\)00029-7](https://doi.org/10.1016/S0747-5632(00)00029-7)
- Capgemini. (2019, September). *Conversational interfaces: The next big thing in customer interaction*. Capgemini. https://www.capgemini.com/wp-content/uploads/2019/09/Report-%E2%80%93-Conversational-Interfaces_Web-Final.pdf
- Corti, K., & Gillespie, A. (2019). A truly human interface: Interacting face-to-face with someone whose words are determined by a computer program. *Journal of Interaction Science*, 7(1), 5. DOI: <https://doi.org/10.3389/fpsyg.2015.00634>
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2019). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*. Advance online publication. DOI: <https://doi.org/10.1007/s11747-019-00696-0>

- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864-886. DOI: <https://doi.org/10.1037/0033-295X.114.4.864>
- Fávero, L. P., & Belfiore, P. (2017). *Manual de análise de dados: Estatística e modelagem multivariada com Excel®, SPSS® e Stata®*. Elsevier Brasil.
- Fan, L., Scheutz, M., Lohani, M., McCoy, M., & Stokes, C. (2017). Do we need emotionally intelligent artificial agents? First results of human perceptions of emotional intelligence in humans compared to robots. In *Intelligent Virtual Agents: 17th International Conference, IVA 2017, Stockholm, Sweden, August 27-30, 2017, Proceedings* (pp. 129-141). Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-67401-8_15
- Gartner. (2023, August 30). Gartner reveals three technologies that will transform customer service and support by 2028. Gartner. <https://www.gartner.com/en/newsroom/press-releases/2023-08-30-gartner-reveals-three-technologies-that-will-transform-customer-service-and-support-by-2028>
- Haas, C., & Moussawi, S. (2020). Are anthropomorphic intelligent agents more intelligent? *AIS Electronic Library (AISel)*. <https://aisel.aisnet.org/>
- Hofacker, C. F., Malthouse, E. C., & Sultan, F. (2016). Big data and consumer behavior: Imminent opportunities. *Journal of Consumer Marketing*, 33(2), 89-97. DOI: <https://doi.org/10.1108/JCM-04-2015-1399>
- Jain, M., Kota, R., Kumar, P., & Patel, S. N. (2018). Convey: Exploring the use of a context view for chatbots. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 468). ACM. DOI: <https://doi.org/10.1145/3173574.3174042>
- Jain, M., Kumar, P., Kota, R., & Patel, S. N. (2018). Evaluating and informing the design of chatbots. In *Proceedings of the 2018 on Designing Interactive Systems Conference 2018* (pp. 895-906). ACM. DOI: <https://doi.org/10.1145/3196709.3196735>
- Jonke, A. W., & Volkwein, J. B. (2018). From tweet to chatbot—Content management as a core competency for the digital evolution. In *Digital Marketplaces Unleashed* (pp. 275-285). Springer. DOI: https://doi.org/10.1007/978-3-662-49275-8_28
- Krasodomski, A., Gwagwa, A., Jackson, B., Jones, E., King, S., Lane, M., ... & Tarkowski, A. (2024). Artificial intelligence and the challenge for global governance. DOI: <https://doi.org/10.55317/9781784136086>
- King, K. (2019). *Using artificial intelligence in marketing: How to harness AI and maintain the competitive edge*. Kogan Page Publishers.
- Klaus, P., & Zaichkowsky, J. (2020). AI voice bots: A services marketing research agenda. *Journal of Services Marketing*, 34(3), 389-398. DOI: <https://doi.org/10.1108/JSM-01-2019-0043>
- Kumar, V., Dixit, A., Javalgi, R. R. G., & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, 44(1), 24-45. DOI: <https://doi.org/10.1007/s11747-015-0426-9>
- Li, J., Monroe, W., Ritter, A., Galley, M., Gao, J., & Jurafsky, D. (2016). Deep reinforcement learning for dialogue generation. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 1192-1202). Association for Computational Linguistics. DOI: <https://doi.org/10.18653/v1/D16-1127>
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*. DOI: <https://doi.org/10.1287/mksc.2019.1192>
- Malhotra, N. K. (2015). *Pesquisa de mercado*. Pearson Ed.
- Marôco, J. (2018). *Análise estatística com o SPSS Statistics: 7ª edição*. ReportNumber, Lda.
- Murphy, M. C., & Dweck, C. S. (2016). Mindsets shape consumer behavior. *Journal of Consumer Psychology*, 26(1), 127-136. DOI: <https://doi.org/10.1016/j.jcps.2015.06.005>
- Oke, A. O., Kamolshotiros, P., Popoola, O. Y., Ajagbe, M. A., & Olujobi, O. J. (2016). Consumer behavior towards decision making and loyalty to particular brands. *International Review of Management and Marketing*, 6(4S), 43-52.
- Shankar, V., Kleijnen, M., Ramanathan, S., Rizley, R., Holland, S., & Morrissey, S. (2016). Mobile shopper marketing: Key issues, current insights, and future research avenues. *Journal of Interactive Marketing*, 34, 37-48. DOI: <https://doi.org/10.1016/j.intmar.2016.03.002>
- Shareef, M. A., Dwivedi, Y. K., & Kumar, V. (2016). Online consumer behavior and marketing. In *Mobile Marketing Channel* (pp. 1-24). Springer. DOI: https://doi.org/10.1007/978-3-319-31287-3_1
- Solomon, M. R., Dahl, D. W., White, K., Zaichkowsky, J. L., & Polegato, R. (2014). *Consumer behavior: Buying, having, and being* (Vol. 10). Pearson.
- Sterne, J. (2017). *Artificial intelligence for marketing: Practical applications*. John Wiley & Sons.
- Tong, S., Luo, X., & Xu, B. (2019). Personalized mobile marketing strategies. *Journal of the Academy of Marketing Science*. DOI: <https://doi.org/10.1007/s11747-019-00693-3>
- Van den Broeck, E., Zarouali, B., & Poels, K. (2019). Chatbot advertising effectiveness: When does the message get through? *Computers in Human Behavior*, 98, 150-157. DOI: <https://doi.org/10.1016/j.chb.2019.04.009>

Wierenga, B., & Van der Lans, R. (Eds.). (2017). *Handbook of marketing decision models* (Vol. 254). Springer.

Zarouali, B., Van den Broeck, E., Walrave, M., & Poels, K. (2018). Predicting consumer responses to a chatbot on Facebook. *Cyberpsychology, Behavior, and Social Networking*, 21(8), 491-497. DOI: <https://doi.org/10.1089/cyber.2017.0518>

Mobile Marketing Association's Brazil. (2019, August 23). Retrieved from: <https://www.mmaglobal.com/research>

Statista. (2024, April 27). Retail mobile commerce revenue worldwide. Retrieved from: <https://www.statista.com/statistics/1449284/retail-mobile-commerce-revenue-worldwide/>

Voicebot.ai. (2021, April 14). U.S. smart speaker growth flat-lined in 2020. Retrieved from: <https://voicebot.ai/2021/04/14/u-s-smart-speaker-growth-flat-lined-in-2020/>