A Work Project, presented as part of the requirements for the Award of a Master's degree in Management from the Nova School of Business and Economics.

CONVERSATIONAL COMMERCE:

ANTHROPOMORPHIC CHATBOTS IN E-COMMERCE AND THEIR EFFECT ON CONSUMER BEHAVIOR

KATHARINA KLEIN

Work project carried out under the supervision of:

Prof. Luis F. Martinez

07-09-2021

Abstract

Conversational agents are becoming increasingly popular in today's technology-driven world, thus a better understanding of factors that enhance customer experience with this technology is crucial. Our study provides insights about the impact of anthropomorphism on consumer behavior in a conversational interface usage scenario. This is the first experimental study to fill the research gap in investigating customer satisfaction with anthropomorphic chatbots in food e-commerce. A sample of 426 participants was tested to verify the proposed hypotheses. The test group interacted with a standard chatbot without human-like characteristics, while the control group communicated with the anthropomorphically designed agent. The results confirm the tremendous potential of anthropomorphic cues in chatbot applications and show that they are positively associated with customer satisfaction and mediated by the variables enjoyment, attitude, and trust.

Key Words: consumer behavior, conversational commerce, e-commerce, chatbot, anthropomorphism, customer experience, customer satisfaction

List of Abbreviations

ADC	Anthropomorphic Design Cues
AI	Artificial Intelligence
AT	Attitude
CA	Conversational Agent
CAGR	Compound Annual Growth Rate
CC	Conversational Commerce
CI	Conversational Interface
CS	Customer Satisfaction
HCI	Human-Computer-Interaction
KMO	Kaiser-Meyer-Olkin Measure
OCE	Online Customer Experience
PE	Perceived Enjoyment
TW	Trust / Trustworthiness
VIF	Variance Inflation Factor

This work was funded by Fundação para a Ciência e a Tecnologia (UID ECO/00124/2019, UIDB/00124/2020 and Social Sciences DataLab, PINFRA/22209/2016), POR Lisboa and POR Norte (Social Sciences DataLab, PINFRA/22209/2016).

Table of Contents

1. Introduction

Who would want to miss out on today's opportunities to interact with others whenever and through whatever channel desired? Especially consumers expect personalized interactions that are accessible at any time and from anywhere. Moreover, such interactions should be tailored to their individual needs and lifestyles. These demands are forcing companies to rethink how they communicate with consumers and differentiate themselves by delivering better customer experiences. The emergence of conversational commerce (CC) represents the next big opportunity for brands and retailers with an expected growing market size of more than 20% per year. The goal of CC is to use conversational interfaces (CIs) to enhance the online shopping customer experience and to satisfy customers' needs in order to increase purchases (Stefanoff, 2021). According to a study by Aspect Consumer Experience Index (2018), more than half of consumers stated interacting with a chatbot application at least once a week and consumer interactions with conversational agents (CAs) will continue to rise. The global CA market is estimated to grow from $4.8bn in 2020 to $13.9bn by 2025 with a CAGR of 21.9% (Marketsandmarkets, 2020). These figures suggest a disruptive transformation of the interaction experience between customers and companies by the integration of CI technologies.

However, research shows that consumers might experience discomfort when they are not convinced of communicating with a human (Luo et al., 2019). Anthropomorphic traits can be added to a CI to simulate a human conversation and influence consumer behavior. Anthropomorphism describes the process of endowing non-human CAs with human features (Epley et al., 2007). Recently, the number of online grocery orders has increased significantly, not least boosted by the Covid-19 pandemic (Repko, 2020). Online food shopping platforms represent a large market in a nearly perfect competitive scenario, which offers new opportunities for participation in the market, especially for start-up entrepreneurs, but also for established retailers (Wang et. al., 2020). This emphasizes the importance for retailers to keep up with the latest technologies and implement CC in their business.

To date, there is limited research on consumers' experiences with anthropomorphic chatbots in food e-commerce, which maintains uncertainty about their impact on consumer behavior. This addresses this gap by exploring the effect of anthropomorphically designed chatbots in an online food retail environment on consumers' perceptions and behavioral decisions. An anthropomorphic chatbot and a baseline version of it as an application on a fictional e-commerce website are contrasted to investigate the impact on consumers' perceived satisfaction, enjoyment, attitude and trust. We thoroughly investigate how users perceive digital agents in CC and what role anthropomorphic cues play in terms of consumers satisfaction. The above objectives are addressed by a research question that summarizes what the present study aims to achieve:

"How does the interaction with anthropomorphic chatbots affect the consumer behavior on an e-commerce platform?"

After introducing CC as well as the concept of Human-Computer-Interaction (HCI), the underlying theory for this paper, anthropomorphism is introduced by means of the prominent SEEK-model of Epley et al. (2007). Scientific insights into anthropomorphically designed CIs, specifically human-like cues in chatbot applications, and their impact on consumer perceptions are given. This is followed by the hypotheses development and data analysis. Using T-test, correlation and mediator analysis, the results are interpreted and implications are given in the discussion section of this paper. The last section provides a conclusion to our key findings.

2. Literature Review

In the following, chatbot technology in e-commerce will be discussed and theories on online customer experience as well as customer satisfaction will be described in more detail. Subsequently, anthropomorphism and its position in HCI, and finally findings on anthropomorphic chatbot design are examined. We mainly refer to sources found in the fields of consumer research, scientific psychology, marketing, and information systems, as well as human-technology interaction and communication.

2.1 Conversational commerce

CC describes the use of chatbots for commercial purposes. Customers can purchase products directly through the digital agent, while the sales process is supported by messages, recommendations, updates, links, or call-to-action buttons. CAs can be programmed to be able to identify purchase intentions and refine offers based on buyers' choices and preferences. They facilitate sales, ordering, and delivery processes for the business and the consumer. The central and primary purpose is to pushing online shop visitors to not just communicate with the business but to ultimately convert them into a customer (Piyush et al., 2016).

"Chatbot" is a hypernym for a CI such as a virtual or digital agent, chatterbot, and conversational agent. The software responds to inputs in natural language and attempts to interact with humans in a voice-based (e.g., Google Cortana) or text-based manner. CAs count to the primary technologies in language-based HCI science. HCI focuses on interaction design and is known for its influence of the connection and communication between machines and human beings (Lazar et al., 2017). Initially, the development of the visual user interfaces was considered the major objective of previous studies conducted in the area of HCI (Følstad & Brandtzaeg, 2017). Today's development of artificial intelligence (AI) and the accessibility to messenger technologies has changed the way humans interact with the devices from the traditional swipe-and-scroll interfaces to conversing in natural language (Etlinger, 2017). AI bots are well suited for companies that need to analyze large amounts of data while learning from the data itself (Joshi, 2020). However, some chatbots have the ability to use simplex techniques for pattern matching and processing strings to engage with consumers, namely rule-based and generative models (Hussain et al., 2019). Unlike AI bots, rule-based ones do not collect data over years to analyze algorithms to understand the consumer. The implementation of the chatbot technology has not yet peaked as the number of businesses adopting it continues to increase (Alger, 2018). In fact, the importance of CAs is growing as our computer technologies continue to evolve and consumer behavior is changing as a result. In the meantime, online users in particular have become accustomed to constantly and quickly interact with their counterparts whenever they want. CAs are often used as an additional option to the regular customer service and provide customers with information for various inquiries. The top priority is always immediateness and comfort (Følstad & Brandtzaeg, 2017). Unlike a physical environment, such as a brick-and-mortar store, customers cannot be screened for purchase intentions, and there is no assessment of the likelihood that an e-store visitor will become a customer. However, CAs can help bring this element to the e-commerce platform (Moe, 2003). Chatbots are about to replace sales assistants with real-time and synchronized two-way interactions, helping companies build relationships with users in the online environment (Yoon et al., 2008).

Online customer experience (OCE) can be considered a major subject for e-retailers in the shopping environment, as the number of touchpoints between customer and company has increased significantly. The customer journey has increased in complexity, and the number of actions to be tracked in a buying process has risen due to the increasing fragmentation of channels (Lemon & Verhoef, 2016). OCE can be described as an individual, multidimensional, psychological reaction to an online platform. Customers perform cognitive and affective conditioning on incoming sensory information from a series of text-based and visual stimuli on a website, which then create an impression in the human's brain (Bleier et al., 2019). According to Bleier et al. (2019), informativeness is the most important cognitive element of OCE that measures the operational component, as well as the importance of the experience to the consumer (Bleier et al., 2019) and the level of comprehensive information on a website (Lim

& Ting, 2012). To improve OCE, Bleier et al. (2019) also mention the importance of entertainment, which is a commitment to the website experience that not only provides a performance-based purchase opportunity but also includes fun and play (Childers et al., 2001; Mathwick et al., 2001). Online marketers should also consider some level of social presence on their online sites, as it can transmit feelings of interpersonal connection through the content (Gefen & Straub, 2003). Consumers might observe human presence and build an emotional connection to a product on a website (Darke et al., 2016). This process is the basis for the induction of higher levels of perceived enjoyment, loyalty and purchase intentions (Wang et al., 2007; Cyr et al., 2007). Finally, a website should also try to stimulate one's senses by visually appealing cues (Schmitt, 1999). Although the online environment may reduce sensory experiences, the senses can be visually stimulated, e.g., by images (Elder et al. 2017). Sensory stimuli have an impact on perceived product performance (Weathers et al., 2007) and purchase behavior (Schlosser, 2003).

Online customer satisfaction (CS) is the result of successful OCE and serves as the key to a company's success and competitiveness (Irfan et al., 2019; Molla & Loicker, 2001). According to Oliver (1997), CS judges the extent to which a service or product satisfies the consumer in a pleasant way. Thus, CS is closely related to customers' affective reactions to the service. Following Oliver's (1997) view, this study understands satisfaction both as a post-consumption evaluation and the overall perceived satisfaction after the interaction with the chatbot on the e-commerce website. Chatbots often have a search or decision support function to create a more convenient, unique, and interactive purchase process. CAs can help to increase CS, build important relationships, reduce uncertainty and anxiety, which enables more efficiency and creates a more comprehensive picture of items and service offerings (Quintino, 2019). Due to their conversation-driven, data-based, and forward-thinking character of chatbots, they play an important role in fostering customer loyalty (Sands et al., 2021). Their main functions are information provision support, navigation assistance for targeted product search, and recommendation provision (Agichtein et al., 2020). Customers can make use of their advantages and receive precise information, obtain guidance and find out about the latest trends in a time-saving manner (Chung et al., 2020). For instance, the chatbot can detect the availability of a particular product and provide information or suggestions about a potential purchase (cross-selling/up-selling). Developers of chatbot applications should pay attention to the accuracy of given recommendations and that they must be convenient, able to adapt, as well as incorporate a level of process efficiency in order to lead to CS (Kraus et al., 2019). Overall, OCE was affirmatively related to CS in e-commerce (Suki & Suki, 2007). Existing studies of new technologies found consumer attitude to be a valuable outcome variable after interacting with a CA (Hassanein & Head, 2005). The influence of enjoyment on technological devices has been literary evaluated several times. Hedonic aspects (e.g., enjoyment) have been revealed to be more important than instrumental properties (e.g., practicability) in e-commerce (Childers et al., 2001). Intensified intrinsic pleasure or joy are factors that generate positive customer attitudes towards online shopping (Kim et al., 2013). Research proves that perceived enjoyment is associated to attitudes on the e-commerce environment (Childers et al., 2001). Araujo (2018) mentioned that CAs with human-like cues are able to affect consumers' attitudes, satisfaction and sensitive attachment to the business and its online appearance. Humanized agents can boost online purchases as they evoke higher levels of perceived empathy and expertise (Luo et al., 2019).

2.2 Anthropomorphism

Anthropomorphism describes the tendency to deploy human attributes, physical or non-physical features, emotions and characteristics to an inhuman object. The main purpose of projecting human-like attributes onto non-human agents is to understand and explain the non-human agents' behavior. Particularly, people with limited time or cognitive resources are more likely to form final judgments that are biased by readily available anthropomorphic knowledge (Epley et al., 2007). The central concept underlying this paper is Epley et al.'s (2007) psychological model of anthropomorphism, which as yet has received limited attention in HCI research. Epley and colleagues' "SEEK" (Sociality, Effectance, Elicited Knowledge) theory helps to explain the practice of anthropomorphism by centralizing factors of the likelihood a person will use anthropomorphism. Firstly, elicited agent knowledge represents the cognitive determinant of anthropomorphism, judging an unfamiliar non-human object. If the item appears to be similar to oneself or to other people, the percipient has a greater likelihood of activating available information about people in order to evaluate it (Epley et al., 2007; Waytz et al., 2010a). Therefore, elicited agent knowledge is strongly influenced by anthropomorphic features (Waytz et al., 2010b). More specifically, the closer the perceptual object approximates a human in regard to observed characteristics and behavior, the more likely people will develop anthropomorphism (Epley et al., 2007). Secondly, sociality motivation indicates the elementary human need to be in social relationships with others. In situations that stimulate the desire for social connection or spawn feelings of loneliness, people tend to anthropomorphize intensified (Epley et al., 2008a; Epley et al., 2008b). Lastly, effectance motivation refers to a humans' fundamental desire to comprehend and maintain command of one's surroundings. Being unaware of a new non-human CA but needing reliance on the CA to perform a particular task, adding anthropomorphism to the CA is especially plausible to reduce perceptions of uncertainty and elevate a sense of trust. In contrast to the underlying cognitive factor, such motivational determinants can best be understood as driven conditions that are induced by a lack of social bond (sociality motivation) or control (efficacy motivation). However, the motivational and cognitive anthropomorphic impacts appear to be unrelated, being based on separate psychological pathways (Epley et al., 2007). In summary, the degree to which someone anthropomorphizes a non-human agent relies on the three determinants described above that alter the activation, correction, or application of anthropomorphic knowledge to a particular target during this inductive reasoning process.

Anthropomorphism has been found to influence consumer behavior, as people tend to feel more engaged and connected to the anthropomorphic object; trustworthiness increases steadily with the extent of anthropomorphism. Some product marketers have already discovered anthropomorphism and are adding human attributes to their goods and services to make them more likable (Aggarwal & McGill, 2007). The literature findings on the impact of anthropomorphism on behavioral intentions can also be confirmed when applied to CAs. One of the desired outcomes of anthropomorphic design in CC is to positively influence human effect, which is reflected as a significant factor observed in HCI. In their research, Nass et al. (1996) were among the first to find that people tend to apply social heuristics to interactions with computers that are interspersed with human or social cues. Social interactions with machines showed an unnatural attribution of anthropomorphic traits to them. This led not only to sociably correct behavior towards the inanimate objects but also to emotional and positive reactions towards computers (Melo et al., 2014). Anthropomorphism in HCI is usually triggered by human-like cues within information technology (IT). Software and hardware developers attempt to use anthropomorphic features to make people feel familiar with the technology in absence of a natural and personal connection. An anthropomorphic software design evokes anthropomorphism, which makes it easier for people to connect with the system and thus facilitates familiarization with its attributes (Burgoon et al., 2000; Epley et al., 2007).

There are several approaches in the literature on how CAs can be anthropomorphized. In this study, we focus on the approach of Go and Sundar (2019), who propose visual cues, identity cues, and conversational cues as humanization tools for chatbots. Human-like visual cues are non-verbal implemented resources that can shape social perceptions through attributes like gestures, pictures or emoticons. Especially emoticons count to the non-verbal symbols transmitting emotive impressions in any text-based and computer mediated communication (Derks et al., 2008). Research on HCI has demonstrated that emoticons contribute to triggering people' s social and emotional reactions (Brown et al., 2016; Wang et al., 2014). Further, the use of human identity cues can easily enhance the agents' humanness. Components such as demographic information or images lead CA users to assess their level of performance depending on their expectations of human agent characteristics (Go & Sundar, 2019). Araujo (2018) attributed the CA a name to address in his experiment and added anthropomorphic conversational cues, which resulted in a stronger anthropomorphism perception for observers than with the non-anthropomorphic version of the object. Conversational cues encompass word or phrase choice and the way in which a narrator describes himself or herself and others (Isbister & Nass, 2000). Conversational cues are able to add anthropomorphism through emotional expression. Studies confirm that varied and context-sensitive responses increase the human-like nature of an agents' verbal behavior (Knijnenburg & Willemsen, 2016; Schuetzler et al., 2014). Anthropomorphism in CA design conveys a sense of efficacy, as the agent's competence then appears magnified to the consumer (Epley et al. 2007). According to Goetz et al. (2003), people rated robots that exhibit playful behavior more extroverted and sociable than serious ones. Consequently, people preferred to cooperate and work with the playful than with the serious robots. Another implication of the effect anthropomorphically designed CAs have on consumers is that they foster the ability to more easily cope with information overload. Lastly, anthropomorphic agents were found to increase consumers' perceived enjoyment and trust, which in turn amplified their intention to use them (Qiu & Benbasat, 2009).

Based on our findings, the first hypothesis represents our general expectation as follows:

> *H1: Consumers interacting with an anthropomorphic chatbot tend to perceive higher levels of anthropomorphic design cues, enjoyment, attitude, trust, and customer satisfaction towards the chatbot, compared to those interacting with a standard chatbot.*

Social presence as a result of anthropomorphic design cues (ADC) is a key dimension of OCE; as CS is defined as a reaction on OCE, we consequently expect a positive relationship between the level of ADC and CS. For this reason, our second hypothesis is as follows:

H2: There is a positive relationship between the level of anthropomorphic design cues and customer satisfaction.

Our findings let us assume, enabling anthropomorphism may serve as an effective method for improving the level of trust (TW), perceived enjoyment (PE), and attitude (AT) interacting with certain technological agents and propose the following hypothesis:

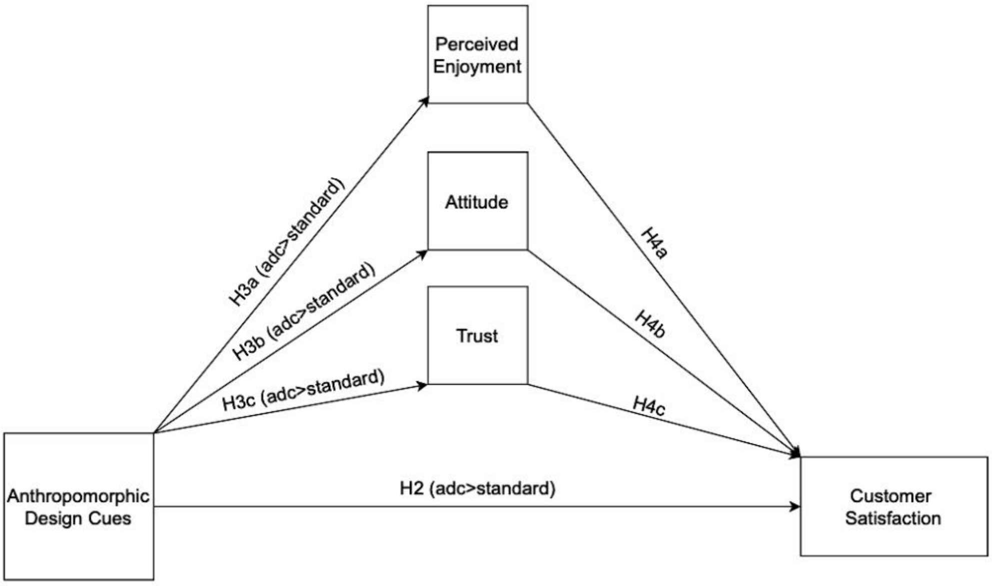
H3: There is a positive relation between anthropomorphic design cues in a chatbot and consumers' a) perceived enjoyment, b) attitude, and c) trust, towards the chatbot.

Osman and Sentosa (2012) found that TW has a mediation effect on CS. Vinerean and Opreana (2014) support this finding and introduce AT as an additional mediator for CS respectively in e-commerce. PE is considered a consequence of successful entertainment. In section 2.1, we presented that entertainment is a key dimension of OCE. Since OCE and CS go hand in hand, we assume that PE can also be considered as a mediator. Following this assumption and our findings in section 2.2, we propose our final hypothesis:

H4: The relation between anthropomorphic design cues and customer satisfaction is mediated by a) perceived enjoyment, b) attitude, and c) trust.

The hypotheses relationships with each variable are depicted in ***Figure 1***. In our conceptual model, ADC is the independent variable and CS the dependent variable, which is mediated by the three mediator variables PE, TW, and AT. The first hypothesis states that all results will be greater for the anthropomorphically designed chatbot (adc) than the standard chatbot (standard) and is noted on the directional arrows for H2 and H3a)b)c).

Figure 1: *Proposed conceptual model*



3. Method

The method briefly describes our research approach including the experimental design, data collection, and sample size. At the end of this section, the derivation and application of the measurement items used within the data collection is explained.

3.1 Research approach

We applied the deductive approach to test for relationships between variables through deducing hypotheses (Saunders et al., 2007) to achieve the research goal. A lack of scientific findings of anthropomorphism in chatbot technology on an e-commerce platform in combination with a potential purchase scenario was identified as the research gap. The few to date existing references about the humanization of CAs focus mainly on chatbots implemented in customer service-related scenarios but not in CC-related ones. The limitation indicated by previous research is data collection via non-experimental approaches only, with the mere description or graphical illustration of an interaction with a chatbot, but not the replication of a real-life scenario on an e-commerce platform (Chung et al., 2020). For testing the proposed hypotheses accordingly, two identical food retail e-commerce platforms were developed with Wix Website Builder. The e-store was specialized in retailing various types of pasta. The pages differed only by the integration of the two different chatbots which were implemented using the Flow XO tool. ***Figure 2*** depicts the visualization of both chatbots; the chatbot system design is attached to ***Appendix 1***.

Figure 2: *Visualization of the research object*

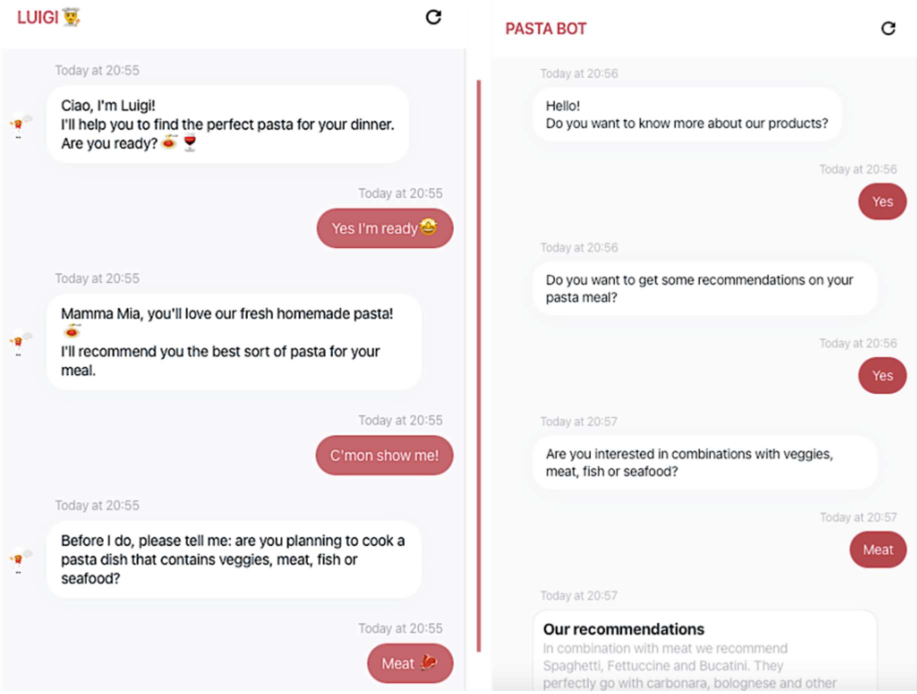


Table 1: *Applied anthropomorphic cues in the chatbot application*

Type	Anthrop. Cue	Example	"Luigi"	Reference
Identity	Visual representation	Images, avatars, faces		Go & Sundar, 2019
Identity	Demographic information	Name, gender, ethnicity	Luigi, male, Italian, chef	Go & Sundar, 2019
Visual	Emoticons	Symbols used to express emotions	🍝🍷🤩👨‍🍳👌😋♡	Wang et al., 2014
Conver sational	Social dialogue	Greeting rituals, anecdotes, non-task related questions	"Ciao, I'm Luigi!“	Bickmore & Picard, 2005;
Conver sational	Emotional expressions	Apologies, congratulations, concerns	"Mamma Mia, you'll love our fresh pasta!"	De Visser et al., 2016
Conver sational	Verbal style	Self-references (“I”) variability of syntax and words	"I'll recommend the best sort of pasta."	Isbister & Nass, 2000;
Conver sational	Temporal cues and reminders	Delayed responses to signal writing; sending reminders	"Just a reminder to answer my question."	Feine et al., 2019

Chatbot 1, hereafter ”Luigi“, was programmed with anthropomorphic features as described in ***Table 1,*** while Chatbot 2, hereafter “standard“, did not show any of these cues, but only a rational identity, visual and conversational design. In section 2.2, the foundation for the development of the anthropomorphic cues shown in the table are outlined. After the conception of the hypotheses model and the creation of the e-commerce platform as well as both chatbots, a structure of factors was established about which to investigate. For testing the hypotheses, research items were added to the variables grounded on theoretical findings as basis for a valid questionnaire (***Appendix 2***).

3.2 Data collection and sample

Primary data collection through an online survey offered the possibility to gather a large amount of records in a short period and is well suited to increase response rates and to be automated. An online questionnaire was primarily spread among German students and contacts through social media as well as on survey exchange networks. The survey was designed with Qualtrics and was active between July 11th, 2021 and August 01st, 2021. It was required to answer five demographic questions designed to confirm sampling criteria, one question about online shopping frequencies, and one about previous chatbot experiences, while 34 questions were asked related to the actual chatbot interaction in the experiment. The survey implied two different scenarios that were randomly assigned to the participants. The candidate was asked to follow either the link to the website where Luigi or the standard chatbot was placed. Arriving on the page, the participant was asked to interact with the respective chatbot and then return to the survey. The non-probability sampling method with convenience and snowball sampling was employed. The survey ran until a reasonable sample size was achieved for each chatbot.

A total of 426 respondents participated in this study; however, after cleaning the raw data, only 371 responses proved significant and were included in our analysis. A distribution balance of 185 respondents exposed to Luigi and, independently to them, 186 people exposed to the standard bot could be obtained. A balance between the two gender types male (n=182) and female (n=189) was attained. Additionally, a third/neutral-gender and an optional choice were given. However, since only 13 participants chose not to specify male or female gender, we decided to consider their response but not to include them in the data analysis. The reason for this is to strive for data homogeneity so that we can subsequently generalize, which is facilitated when a large number of people choose an option. In detail, this affected two responses for ”non-binary/third gender“, seven for ”prefer not to say“, and four responses left blank. The age structure of the respondents was divided into the four main popular generations; Gen Z, Y, X and baby boomers (Francis & Hoefel, 2018). We did not expect participants from another age category than the presented. In fact, 65% of the attendees were between 25 and 40 years old (*n*=251), 31% chose the age category 14 to 24 (*n*=120), while the remaining 4% counts for people older than 40 years (*n*=16). Due to the low response rate from generation X and baby boomers and the desire for homogeneity and significance, only generation Z and Y were included in our analysis. Most of the participants were academics; the underlying cause to this is that the presented study was ultimately conducted as part of a master’s thesis, by convenience and snowball sampling. Despite a broad mass of German respondents, even other nationalities across Europe, the USA, Africa, India, and Australia were represented. The sample details of the above described demographic data, as well as the number of different educational groups of our cleaned raw data, can be found in ***Appendix 3***. To get a better picture of their skills in using the online shopping environment, candidates also provided information about shopping frequencies and indicated whether they had any prior touchpoints with chatbots (***Appendix 4***). The most important information we can extract from this data is that 86% of respondents ever had consciously contact with a chatbot, and most of the respondents usually doing online purchases more than once a month (36%).

3.3 Measurement

After the data collection of demographics and online shopping habits in the first part of the survey, five constructs were measured after the interaction with the chatbot had taken place. The first one to be assessed was *anthropomorphic design cues (ADC)* using nine items that measured social presence and anthropomorphism, as discussed by Go and Sundar (2019), Nowak and Rauh (2005) and Goetz et al. (2003). When the context required it, a content adjustment was made to change the wording from “avatar” or “human” to “chatbot”. C*ustomer satisfaction (CS)* could be captured using Chung et al.’s (2020) approach. Their scales were appropriate for this study because they refer to a similar scenario in their research. CS was measured, e.g., whether expectations are met or if the chatbot did a good job. Respondents’ *perceived enjoyment (PE)* of the chatbot was measured by six items used by Mikalef et al. (2013) testing for hedonic motivation. Zarantonello and Pauwels-Delassus (2017) suggest in their manual that for assessing *trust or trustworthiness (TW),* the scale dimensions can be divided into competence and benevolence. In addition, trust can also be used to measure the relationship with a project or a brand; we have taken items from all three dimensions and adapted them to chatbots. Venkatesh et al. (2003) describe items related to the user acceptance and *attitude (AT)* towards IT. We were able to adopt three of them to our measurement model, as the chatbot is defined as a technology. Spears and Singh (2004) published an entire paper on conceivable variables on attitude out of which we applied two more pertinent scales. As the number of scale points increases, the information content of the corresponding items becomes more refined. In contrast to an even number of scale points, the respondent can choose a neutral position for an odd number of response alternatives and does not have to decide on an inclination (Lehmann et al., 1993). For this reason, a 7-point Likert scale (from “strongly disagree” to “strongly agree”) is used to assess the preferences. ***Appendix 2*** shows the variables, their definition, and the respective measurement items based on the designated literature.

4. Results

The collected data was analyzed using SPSS software version 27. Single missing values could be replaced using single imputation in the data analysis software. After cleaning the data, both descriptive and inferential statistics were applied to exemplify the sample. First, the conceptual model was tested on validity and reliability by performing confirmatory factor analysis and calculating Cronbach's alpha coefficient. An independent samples t-test was conducted to assess the influence of the anthropomorphic and standard chatbot on CS. The correlation analysis was conducted to determine the effect size of the relationship between the respective variables. Lastly, regressions were calculated using Hayes' (2018) mediator model to measure indirect effects between the variables.

4.1 Validity and reliability testing

Outliers could have been detected with the help of boxplots and were eliminated, resulting in a normal distribution of the dataset with its scales in the range of an Asymp. sig. (2-tailed) $p > 0.05$ measured by the Kolmogorov-Smirnov test. In order to test the data fit and validness of the measurement model, a confirmatory factor analysis was deployed. Both Bartlett's measure as well as the Kaiser-Meyer-Olkin test (KMO) of Sampling Adequacy indicate that the variables are suitable for factor analysis. Evidence for this is a chi-square range in Bartlett's test between a minimum of 10 for CS and a maximum of 36 for ADC, as well as $p < .001$. The KMO value with a minimum of .892 is significantly above the recommended value (Field, 2013), which means that a principal component analysis can be performed; the exact values are represented in ***Appendix 5.*** The merging of research items belonging to one variable that originated from two different literature sources (the case with ADC and AT) turned out to be reasonable due to the positive results for the factor analysis. Afterwards, we conducted the output validity test using Pearson Product Moment Correlations to determine whether the questionnaire was valid. Every item could be verified valid as we obtained a Sig. (2-tailed) of $p < .001$. In the next step, we used Cronbach's alpha as our measurement for expressing the internal consistency of the data collection instrument. The results indicate a total reliability coefficient of .992 for all 34 items. They range between .850 and .926 for the single items, which reveals high reliability to predict the variable (Hair et al., 2010). ***Appendix 6*** resumes the reliability measurement for every scale to each chatbot.

4.2 T-test statistics

Independent samples t-tests were applied in order to investigate the significance of the differences between the groups (male vs. female; generation Z vs. generation Y; anthropomorphic chatbot vs. standard chatbot). The mean values let us undertake that males consistently had a higher tendency to agree with all statements ($M_{average}$= 4.3, $SD_{average}$= 1.7) than females ($M_{average}$= 4.1, $SD_{average}$= 1.7). However, taking Levene's test for equality of variances into account, we were able to determine that the descriptively examined mean values are not statistically significant and we indeed have to neglect gender differences for every variable. Regarding age, Levene's tested $p > .05$ for the variables ADC, PE, and TW, which means that age influences only those variables, while CS and AT are unaffected. The independent samples t-test was especially helpful to validate that both chatbots are perceived significantly different by the consumers. In terms of ADC, we found higher perceived levels of anthropomorphism interacting with Luigi (M= 5.36; 7-point Likert scale) than with the standard bot (M= 2.68; 7-point Likert scale), $t(367.011) = 23.683$, $p < .05$. Hence, an important precondition is fulfilled as the participants perceive the anthropomorphic chatbot as such. Further, all scores for the standard chatbot are significantly lower than for Luigi. Based on those results, we are able to confirm our first hypothesis H1. ***Appendix 7*** provides an overview of the mean values conducted with the independent samples t-test on a confidence level of .95.

4.3 Correlation analysis

A Pearson correlation analysis for both samples was performed to determine the effect size between the dependent and independent variables and to receive anticipations about the validity of H2 as well as H3a, H3b, and H3c. The second hypothesis suggests the existence of a positive relationship between the level of ADC and CS; H3 expects the level of ADC to positively influence PE, AT and TW. The results indicate a significant and positive relation between ADC and CS (r= .819, p < .001) for Luigi and (r= .616, p < .001) for the standard bot. As the variables are highly correlated we can statistically support H2. Equally, the relation between ADC and PE (r_{Luigi}= .800; $r_{Standard}$= .613; p < .001), AT (r_{Luigi}= .625; $r_{Standard}$= .468; p < .001) and TW (r_{Luigi}= .728; $r_{Standard}$= .505; p < .001), present us significant correlations. Hence, H3a, H3b, and H3c are also validated. However, coefficients r > .5 indicate a high correlation, which occurred for all variables except the relationship between ADC and AT for the standard bot; although a moderately significant correlation is given here. The exact values can be found in the correlation tables attached to ***Appendix 8***. Testing for multicollinearity, the VIF value of the coefficients were analyzed. Since all values lie between 1-10, there is no indication for multicollinearity.

4.4 Mediation analysis

We conducted a mediation analysis to test the effect of ADC (X) on the outcome variable CS (Y), adding the mediators PE, AT and TW (M) to the model as illustrated in ***Appendix 9***. The analysis was completed with the PROCESS macro by Hayes (2018), which processes ordinary least squares regression, yielding unstandardized path coefficients for total, direct and indirect effects. Bootstrapping with 5000 samples together with heteroscedasticity consistent standard errors (MacKinnon, 2007) were used to calculate confidence intervals and inferential statistics. Effects were deemed significant when the confidence interval did not include zero. The detailed matrix can be retrieved in ***Appendix 10.***

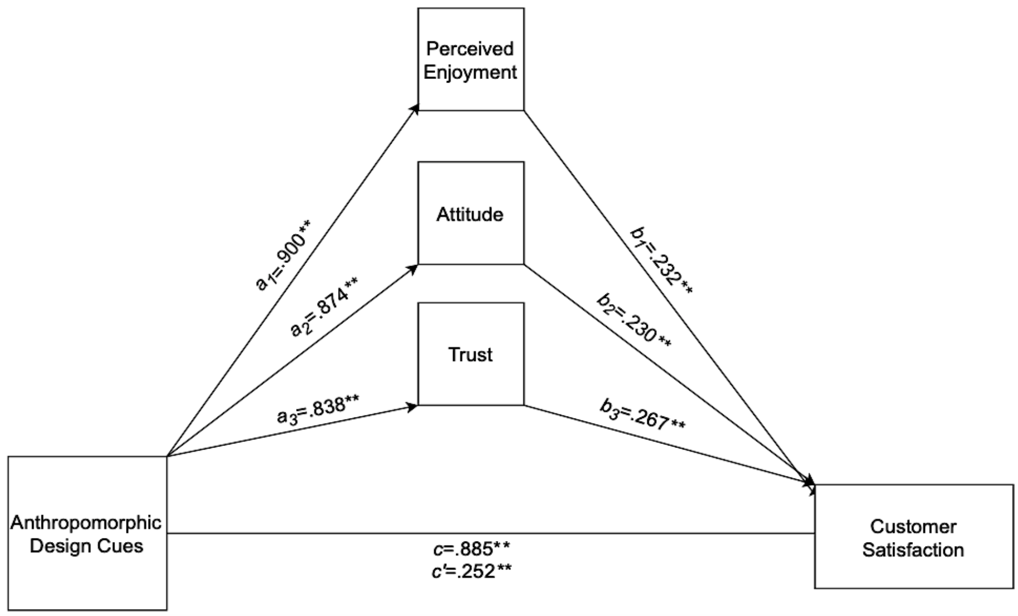
Mediation occurs when the direct part coefficient between the independent and dependent variable decreases as soon as the indirect path through the mediator is established in the model. According to Baron and Kenny (1986), all conditions for mediation are met: Firstly, the direct path between ADC (X) and CS (Y) was assessed without the intervention of the mediators. The direct path coefficient (c) was b= .885, p < .001 and then changed after the introduction of the mediators (PE, AT, TW) to *b*= .252, *p* < .001 (c'). The amount of the decrease of the relationship between X and Y accounted by M is .663 which represents 75% of the total effect. After adding the mediators, the predictor variable (ADC) predicted the outcome variable (CS) less strongly; ADC significantly predicted the mediators (path a: b_{PE}= .900; b_{AT}= .874, b_{TW}= .839; *p* < .001), which, in turn, significantly predicted CS (path b: b_{PE}= .232, b_{AT}= .229, b_{TW}= .267; *p* < .001). The relationship between ADC and CS is mediated by every of the mediator variables with an indirect effect *ab*= .633, 95% - CI[.535, .722]. The three variables are considered as partial mediators as X (ADC) significantly and directly effects Y (CS). Based on the results we can confirm our last hypothesis H4. ***Table 2*** summarizes the results at a glance; ***Figure 3*** depicts them on the model.

Table 2: *Bootstrapping results of the mediation model*

Path	*Coefficient b*	df	t-value	p-value	LLCI[a)]	ULCI[b)]
a1 *(ADC → PE)*	.900	368	37.277	.000	.8527	.9477
a2 *(ADC → AT)*	.874	365	33.276	.000	.8223	.9256
a3 *(ADC → TW)*	.838	368	30.468	.000	.7846	.8928
b1 *(PE → CS)*	.232	365	3.891	.000	.1148	.3493
b2 *(AT → CS)*	.230	365	4.880	.000	.1370	.3218
b3 *(TW → CS)*	.267	365	5.072	.000	.1634	.3704
c *(ADC → CS)*	.885	368	36.972	.000	.8383	.9325

Notes. a) LL = Lower Level; b) UL = Upper Level; Confidence level for all CIs = 95.0%; Number of bootstrap samples = 5000;

Figure 3: *Applied mediation model with regression coefficients*



Note: ** p< .001

5. Discussion

Our study aimed to examine the factors that have an impact on CS in 2 different CC usage scenarios. In particular, the effect of anthropomorphic chatbots on consumer behavior has been investigated. The following section focuses on the main findings and explains how they relate to the initial literature review and the respective research question. Consumer behavior is the study of why and how people consume products and services. Consumers' behavior can broadly be attributed to three main influences – the characteristics, environment, and genetics of the individual (Chaudhuri, 2006). We can influence peoples' behavioral response by specific stimuli – in our study we used ADC as a stimulus for effecting CS in CC.

Contrary to our expectations, we found that there is no significant effect of the gender variable on our model, even though a slight difference could be identified when analyzing the mean values. Men could have been expected to rate the scales higher as they might be more aware of technology and show a greater interest in chatbot applications. However, the t-test for the age groups found that age influences at least three variables. We suppose that the younger the interactors are, the more the chatbot might appear "common" to the people interacting with it, as they grew up with the latest technologies. Our collected data about previous chatbot interaction and online shopping habits reflect the importance of engagement through CIs and is thereby in line with our research about the growth opportunities on the global CA market (Marketsandmarkets, 2020). We revisited people's scales according to their online shopping behavior and previous chatbot interaction. Those who purchase online only 1-3 times a year rated scales for all variables the lowest. This result makes us interpret that those people do not enjoy shopping online or only do so when necessary. In summary, the development of chatbot features is most important for the engagement of sophisticated online shoppers, rather than those who rarely buy anyway. Worth mentioning as well, people who had never interacted with a chatbot before perceived significantly fewer anthropomorphic cues than people who had previous chatbot experience. This is not entirely consistent with the SEEK theory (Epley et al., 2007), whose paper asserts that anthropomorphizing is more likely to occur when interacting with unknown non-human objects. Conversely, Epley et al. (2007) did not demonstrate whether individuals who have already been exposed to a similar application anthropomorphize more than individuals who have not. However, the elicited agent theory, which states that an object to which anthropomorphic properties are added is more likely to be anthropomorphized coincides with our observations since we received significantly higher ADC values for Luigi compared to the standard bot.

Based on our findings, we claimed that the variables PE, AT, TW and CS will be rated higher when interacting with an anthropomorphic CA. Indeed, we could prove higher levels of those variables in the anthropomorphically designed chatbot compared to the standard bot scenario. Thus, our predictions covered in H1 could be validated. Our assumption that a positive relationship between ADC and CS prevails (H2) is primarily based on the conclusion that ADC leads to higher levels of PE, AT and TW (H3), which are considered essential for eliciting OCE, and CS as logical consequence of OCE. Our statistical evidence confirmed that we correctly proposed a significant relationship between ADC and CS (H2), congruently with Luo et al. (2019), who predict higher CS when communicating with humanized agents. Further, and not otherwise expected due to the H2 testing results, a correlation analysis confirmed the positive relation of ADC with each of the three variables PE, AT and TW (H3). However, our mediator analysis validated likewise that these three variables are caused by ADC and ultimately reinforce higher CS (H4). In detail, PE, as well as AT and TW could be considered as significant partial mediators between ADC and CS. This implies that PE, AT and TW explain the relationship between ADC and CS. More precisely, ADC leads to higher levels of CS when the level of PE, AT and TW is high. This partial mediation is consistent with Baron and Kenny

(1986), who claim that partial mediation is common in social science research as full mediation would be unrealistic.

5.1 Theoretical implications

The present study enriches the literature at experimental research on an innovative marketing application by analyzing the variables that play an important role for a successful chatbot interaction within an online purchase scenario. After a comprehensive review of the existing literature on the incorporation of chatbots in e-commerce, a research gap was discovered, particularly with respect to experimental studies on anthropomorphism in CC. We demonstrated that the construct of anthropomorphism can also be applied to CA technologies (Burgoon et al., 2000) and supplement the literature on the anthropomorphism theory associated with CC, showing that a chatbot embedding human-like cues is capable of generating better OCE and, in turn, higher CS (Hassanein & Head, 2007; Epley et al. 2007). Our results provide knowledge about anthropomorphic visual and conversational chatbot design, and its impact on perceived humanness. The mediation analysis allowed us to identify mediators involved in the generation of CS. We have continued theories of Vinerean and Opreana (2014) who discovered AT as a mediator to CS, and Osman and Sentosa (2012) who noted TW as a mediator in e-commerce. Derived from our inferences, we extended existing theories, introducing PE as a mediator into our model and proved that this variable amplifies CS. This paper aims to highlight the importance of appropriate chatbot design when seeking higher levels of PE, TW and AT throughout a CA interaction; and that these very variables should be considered in order to satisfy customers in the context of CC.

5.2 Practical implications

Our results could help marketers with their decisions about the use and design of conversational tools on their online platforms. Not only did we prove multiple times that chatbot technology is an effective way to reach customers, but also how the application needs to be designed to attract them. Specifically, this study aims to encourage retailers to use chatbots as a sales channel (Yoon et al., 2008). Chatbots used in retail should lead to enjoyment, inspire trust and generate a good attitude towards the application. In this way, customers can be attracted to the e-store through pleasant experiences. Retailers that want to create positive consumer impressions should be empathetic and build both a lasting social bond and engage with conversations that include small talk, sympathetic feedback, emoticons or, images to create anthropomorphism and consequently increase CS (Go & Sundar, 2019). Higher levels of CS usually lead to the retailers' main goal - enhanced purchase intentions (Luo et al., 2019).

5.3 Limitations and future research

Despite the valuable findings and insights gained, this study as part of a master‘s thesis in consumer behavior was limited in time and resources. Thus, some limitations should be considered when interpreting the results and conclusions. Firstly, the scientific background of this experimental study and the short time frame led to technical limits. It was not possible to develop a mature chatbot technology for the subject of this study, but rather to resort to a minimum viable application, more precisely a rule-based chatbot with predefined answers. The CC scenario was recognizable as an experiment and may have caused some bias among respondents. The e-store offered only a few products without providing a full checkout process, as it stopped after the shopping cart was filled and did not move to the checkout process. Our recommendation for future studies is to carry out experiments with the help of IT experts and perform them in a real-life scenario, on a proper e-commerce platform. An investigation in cooperation with global food retailers, such as Lidl or Aldi, would exceptionally be valuable. Further, studies could address the technological progress, as it seems to be a promising approach to investigate AI developments in the future. Within our study, we could not integrate AI, as a longer time frame would have been required to learn artificially. However, implementing AI in experimental studies could enrich literature on HCI. There were also limitations within our proposed conceptual model. We examined CS as an important variable in consumer behavior, yet marketers should also consider purchase intentions in their strategies. Due to the minimum viable test scenario, we decided not to measure this variable in our model. Therefore, future research could investigate the impact of anthropomorphic CAs on consumer purchase intention in an enhanced or real-life scenario. Lastly, well educated individuals from generations Z and Y predominated our sample. Sampling methods other than non-probability and snowballing could be applied by upcoming studies to get insights on customer experiences across all age groups and education levels.

6. Conclusion

In this study, we have outlined the importance of anthropomorphism in the context of CC, which could be beneficial to both businesses and consumers. Anthropomorphism, which can attribute human characteristics to technology, can be a useful tool for improving communication and enhancing consumer trust and empathy. Retailers need to assess the appropriate combination of conversational and visual representation of their CIs to match the context and their brand identity. When implementing chatbot technologies on an e-commerce platform, attention must be paid to the appropriate environment, the desired accomplishment, the way information is presented, as well as user feedback, and offered user choices. A chatbot interacting with many consumers might entail higher consequences on brand perception than a conversation with a retail salesperson. CC must not be designed and implemented with technology alone in mind, but requires a far more nuanced approach, e.g., tech companies like Google and Amazon already have taken steps to eliminate gender bias in conversational design (Specia, 2019). Identifying appropriate opportunities, incorporating the human touch, and navigating a growing list of security, ethical, and moral tensions cannot be ignored. CC as an innovative research area has a great future perspective, yet some investigation gaps will need to be addressed in upcoming studies.

7 References

Aggarwal, P., & McGill, A. (2007). Is That Car Smiling at Me? Schema Congruity as a Basis for Evaluating Anthropomorphized Products. *Journal of Consumer Research, 34*(4), 468–479.

Agichtein, E., Hakkani-Tur, D., Kallumadi, S., & Malmasi, S. (2020). ConvERSe’20: the WSDM 2020 workshop on conversational systems for e-commerce recommendations and search”. *Proceedings of the 13th International Conference on Web Search and Data Mining*, 897-898.

Alger, K. (2018). *Forbes*. Retrieved June 9, 2021, from https://www.forbes.com/sites/delltechnologies/2018/12/13/to-bot-or-not-the-rise-of-ai-chatbots-in-business/?sh=712bd934375d

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.

Aspect. (2018). *Aspect Software Inc.* Retrieved May 21, 2021, from Aspect Consumer Experience Index Survey.: https://www.aspect.com/globalassets/files/ebooks/2018_Aspect_Consumer_Experience_Index_Results.pdf

Baron, R., & Kenny, D. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology 51*(6), 1173-1182.

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction, 12*(2), 293-327.

Bleier, A., Harmeling, C. M., & Palmatier, R. (2019). Creating Effective Online Customer Experiences. *Journal of Marketing, 83*(2), 98-119.

Brown, S. A., Fuller, R., & Thatcher, S. M. (2016). Impression Formation and Durability in Mediated Communication. *Journal of the Association for Information Systems, 17*(9), 614-647.

Burgoon, J., Bonito, J., 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.

Chaudhuri, A. (2006). *Emotion and Reason in Consumer Behavior.* Burlington: Routledge.

Chen, J.-S., Le, T.-T., & Florence, D. (2021). Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing. *International Journal of Retail & Distribution Management, Vol.*(No.) ahead-of-print, pp-pp ahead-of-print.

Childers, T. L., Carr, C. L., Peck, J., & Carson, S. (2001). Hedonic and Utilitarian Motivations for Online Retail Shopping Behavior,. *Journal of Retailing, 77*(4), 511–35. Chung, M., Ko, E., Joung, H., & Kim, S. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research, 117*(9), 587-595.

Cyr, D., Hassanein, K., Head, M., & Ivanov, A. (2007). The Role of Social Presence in Establishing Loyalty in E-Service Environments. *Interacting with Computers, 19*(1), 43-56.

Darke, P. R., Brady, M., Benedicktus, R. L., & Wilson, A. E. (2016). Feeling Close from Afar: The Role of Psychological Distance in Offsetting Distrust in Unfamiliar Online Retailers. *Journal of Retailing, 92*(3), 287–99.

de Visser, E. J., Monfort, S. S., McKendrick, R., Smith, M. A., McKnight, P. E., Krueger, F., & Parasuraman, R. (2016). Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology, 22*(3), 331-349.

Derks, D., Bos, A. E., & Grumbkow, v. J. (2008). Emoticons in Computer-Mediated Communication: Social Motives and Social Context. *Cyber Psychology & Behavior, 11*(1),99-101.
Elder, R. S., Schlosser, A. E., Poor, M., & Lidan, X. (2017). So Close I Can Almost Sense It: The Interplay Between Sensory Imagery and Psychological Distance. *Journal of Consumer Research, 44*(4), 877–94.
Epley, N., Akalis, S., Waytz, A., & Cacioppo, J. T. (2008a). Creating Social Connection Through Inferential Reproduction. *Psychological Science, 19*(2), 114-120.
Epley, N., Waytz, A., & Cacioppo, J. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review, 114*(4), 864-886.
Epley, N., Waytz, A., Akalis, S., & Cacioppo, J. T. (2008b). When We Need A Human: Motivational Determinants of Anthropomorphism. *Social Cognition, 26*(2), 143-155.
Etlinger, S. (2017). *The Conversational Business: How Chatbots Will Reshape Digital Experiences.* Retrieved June 12, 2021, from Altimetergroup: https://andyblackassociates.co.uk/wp-content/uploads/2015/06/ConversationalBusiness_FINAL-1.pdf
Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. *Journal of Human Computer Studies, 132*, 138-161.
Field, A. (2009). Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll). Los Angeles: Thousand Oaks, Calif.: SAGE Publications.
Følstad, A., & Brandtzaeg, P. (2017). Chatbots and the new world of HCI . *Interactions, 24*(4), 38-42.
Francis, T., & Hoefel, F. (2018). *McKinsey*. Retrieved July 31, 2021, from ‘True Gen’: Generation Z and its implications for companies: https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/true-gen-generation-z-and-its-implications-for-companies#
Gefen, D., & Straub, D. (2003). Managing User Trust in B2C E-Services. *E-Service Journal*, *2*(2), 7-24.
Go, E., & Sundar, S. (2019). Humanizing Chatbots: the effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior, 97,* 304-316.
Goetz, J., Kiesler, S., & Powers, A. (2003). Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation. *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication*, 55-60.
Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate Data Analysis, Vol.*7. New Jersey.: Prentice Hall, Upper Saddle River.
Hassanein, K., & Head, M. (2005). The impact of infusing social presence in the web interface: an investigation across product types. *International Journal of Electronic Commerce*, *10*(2), 31-55.
Hayes, A. F. (2018). Partial, conditional, and moderated mediation: Quantification, inference, and interpretation. *Communication Monographs, 85*(1), 4–40.
Hussain, S., Sianaki, O., & Ababneh, N. (2019). A Survey on Conversational Agents/Chatbots Classification and Design Techniques. *Primate Life Histories, Sex Roles, and Adaptability*, *927*, 946-956.
Irfan, W., Siddiqui, D., & Ahmed, W. (2019). Managing in-store logistics: perspective from Pakistani small and medium retail stores. *International Journal of Retail and Distribution Management, 47*(4), 350-367.
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.

Joshi, N. (2020). *Forbes*. Retrieved June 12, 2021, from https://www.forbes.com/sites/cognitiveworld/2020/02/23/choosing-between-rule-based-bots-and-ai-bots/

Kim, H., Suh, K., & Lee, U. (2013). Effects of collaborative online shopping on shopping experience through social and relational perspectives. *Information and Management, 50*(4), 169-180.

Knijnenburg, B. P., & Willemsen, M. C. (2016). Inferring Capabilities of Intelligent Agents from Their External Traits. *ACM Transactions on Interactive Intelligent Systems, 6*(4), 28-25.

Kraus, D., Reibenspiess, V., & Eckhardt, A. (2019). How Voice Can Change Customer Satisfaction: A Comparative Analysis between E-Commerce and Voice Commerce. *Internationale Tagung Wirtschaftsinformatik, Siegen*, *14*, 23-27.

Lazar, J., Hochheister, H., & Feng, H. (2017). *Research Methods in Human-Computer Interaction, Vol. 2*. Cambridge: Todd Green.

Lee, J., Pi, S., Kwok, R., & Huynh, M. (2003). The contribution of commitment value in Internet commerce: an empirical investigation. *Journal of the Association for Information Systems, 4*(1), 2.

Lehmann, E. L. (1993). The Fisher, Neyman–Pearson theories of testing hypotheses: One theory or two? *Journal of the American Statistical Association, 88*(424), 1242–1249.

Lemon, K., & Verhoef, P. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing, 80*(6), 69-96.

Lim, W., & Ting, D. (2012). E-shopping: An Analysis of the Uses and Gratifications Theory. *Modern Applied Science, 6*(5), 48-63.

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 38(6)*, 937-947.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual review of psychology*, *58*(1), 593-614.

Marketsandmarkets. (2019). *Conversational AI Market by Component, Type, Technology, Application, Deployment Mode, Vertical, and Region - Global Forecast to 2024*. Retrieved May 10, 2021, from https://www.marketsandmarkets.com/Market-Reports/conversational-ai-market- 49043506.html

Mathwick, C., Naresh, M., & Rigdon, E. (2001). Experiential Value: Conceptualization, Measurement and Application in the Catalog and Internet Shopping Environment. *Journal of Retailing*, *77*(1), 39–56.

Melo de, C., Carnevale, P., Read, S., & Gratch, J. (2014). Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology*, *106*(1), 73-88.

Mikalef, P., Giannakos, M., & Pateli, A. (2013). Shopping and word-of-mouth intentions on social media. *Journal of Theoretical and Applied Electronic Commerce Research, 8*(1), 5-34.

Moe, W. (2013). Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream. *Journal of Consumer Psychology*, *13*(1-2), 29-39.

Molla, A., & Licker, P. (2001). E-commerce systems success: an attempt to extend and respecify the Delone and MacLean model of IS success. *Journal of Electronic Commerce Research*, *2*(4), 131-141.

Nass, C., Fogg, B., & Moon, Y. (1996). Can Computers be Teammates? *International Journal of Human-Computer Studies, 45*(6), 669–678.

Nowak, K. L., & Rauh, C. (2005). The Influence of the Avatar on Online Perceptions of Anthropomorphism, Androgyny, Credibility, Homophily, and Attraction. *Journal of Computer-Mediated Communication, 11*(1), 153-178.

Oliver, R. (1997). *Satisfaction: A Behavioral Perspective on the Consumer.* Boston, MA.: Irwin McGraw-Hill.

Opreana, A., & Vinerean, S. (2014). Analyzing Mediators of the Customer Satisfaction – Loyalty Relation in Internet Retailing. *Expert Journal of Marketing, 2*(1), 1-14.

Osman, Z., & Sentosa, I. (2013). A study of mediating effect of trust on customer satisfaction and customer loyalty relationship in Malaysian rural tourism. *European Journal of Tourism Research, 6*(2), 192-206.

Piyush, N., Choudhury, T., & Kumar, P. (2016). Conversational commerce a new era of e-business. *International Conference System Modeling & Advancement in Research Trends (SMART).* IEE 2016, 322-327.

Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, *25*(4), 145-182.

Repko, M. (2020). *CNBC*. Retrieved June 2, 2021, from Grocers limit food purchases, urge shoppers not to hoard as panic buying continues. Kroger ramps up hiring: https://www.cnbc.com/2020/03/14/grocers-limit-food-purchases-urge-shoppers-not-to-hoard-kroger-is-hiring.html

Rose, S., Clark, M., Samouel, P., & Hair, N. (2012). Online Customer Experience in E-Retailing: An Empirical Model of Antecedents and Outcomes. *Journal of Retailing, 88*(2), 22-308.

Sands, S., Ferraro, C., Campbell, C., & Tsao, H.-Y. (2021). Managing the human–chatbot divide: how service scripts influence service experience. *Journal of Service Management, 32*(2), 246-264.

Saunders, M., Lewis, P., & Thornhill, A. (2007). *Research Methods For Business Students* (Vol. 4). Harlow: Pitman Publishing.

Schlosser, A. E. (2003). Experiencing Products in the Virtual World: The Role of Goal and Imagery in Influencing Attitudes Versus Purchase Intentions. *Journal of Consumer Research, 30*(2), 98-184.

Schmitt, B. (1999). Experiential Marketing. *Journal of Marketing Management, 15*(1/3), 53–67.

Schuetzler, R. M., Grimes, M., Giboney, J. S., & Buckman, J. (2014). Facilitating natural conversational agent interactions: lessons from a deception experiment. *In Proceedings of the Thirty Fifth International Conference on Information Systems, Auckland, 16,* 1-16.

Spears, N., & Singh, S. (2004). Measuring Attitude toward the Brand and Purchase Intentions. *Journal of Current Issues & Research in Advertising, 26*(2), 53-66.

Specia, M. (2019). *NY Times*. Retrieved August 6, 2021, from Siri and Alexa Reinforce Gender Bias, U.N. Finds: https://www.nytimes.com/2019/05/22/world/siri-alexa-ai-gender-bias.html

Stefanoff, S. (2021). *Deloitte*. Retrieved May 21, 2021, from https://www2.deloitte.com/au/en/blog/consumer-blog/2020/unlocking-potential-of-conversational-commerce-in-retail.html

Suki, N., & Suki, N. (2007). Online buying innovativeness: effects of perceived value, perceived risk and perceived enjoyment. *International Journal of Business and Society, 8*(2), 81-93.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS quarterly, 27*(3), 425-478.

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.

Wang, L. C., Baker, J., Wagner, J. A., & Wakefield, K. (2007). Can a Retail Web Site Be Social? *Journal of Marketing*, *71*(3), 57-143.

Wang, W., Zhao, Y., Qiu, L., & Zhu, Y. (2014). Effects of Emoticons on the Acceptance of Negative Feedback in Computer-Mediated Communication. *Journal of the Association for Information Systems,, 15*(8), 454-483.

Wang, Y., Xu, R., Schwartz, M., Ghosh, D., & Chen, X. (2020). COVID-19 and Retail Grocery Management: Insights from A Broad-based Consumer Survey. *IEEE Engineering Management Review, 48*(3), 202-211.

Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010b). Causes and consequences of mind perception. *Trends in Cognitive Sciences, 14*(8), 383-388.

Waytz, A., Morewedge, C. K., Epley, N., Monteleone, G., Gao, J.-H., & Cacioppo, J. T. (2010a). Making sense by making sentient: Effectance motivation increases anthropomorphism. *Journal of Personality and Social Psychology, 99*(3), 410-435.

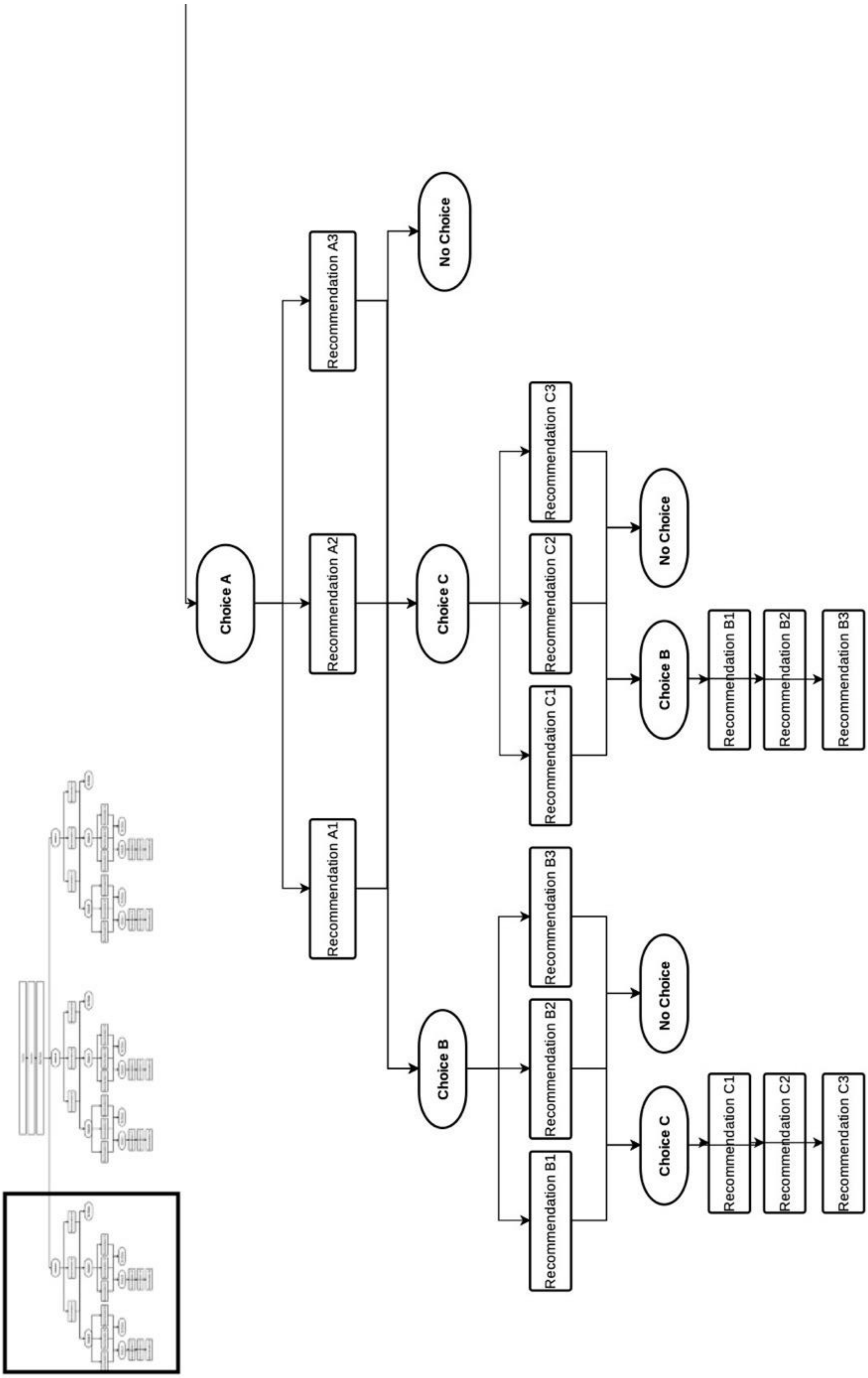
Weathers, D., Sharma, S., & Wood, S. L. (2007). Effects of Online Communication Practices on Consumer Perceptions of Performance Uncertainty for Search and Experience Goods. *Journal of Retailing*, *83*(4), 393–401.

Yoon, D., Choi, S., & Sohn, D. (2008). Building customer relationships in an electronic age: The role of interactivity of E-commerce Web sites. *Psychology and Marketing, 25*(7), 602-608.

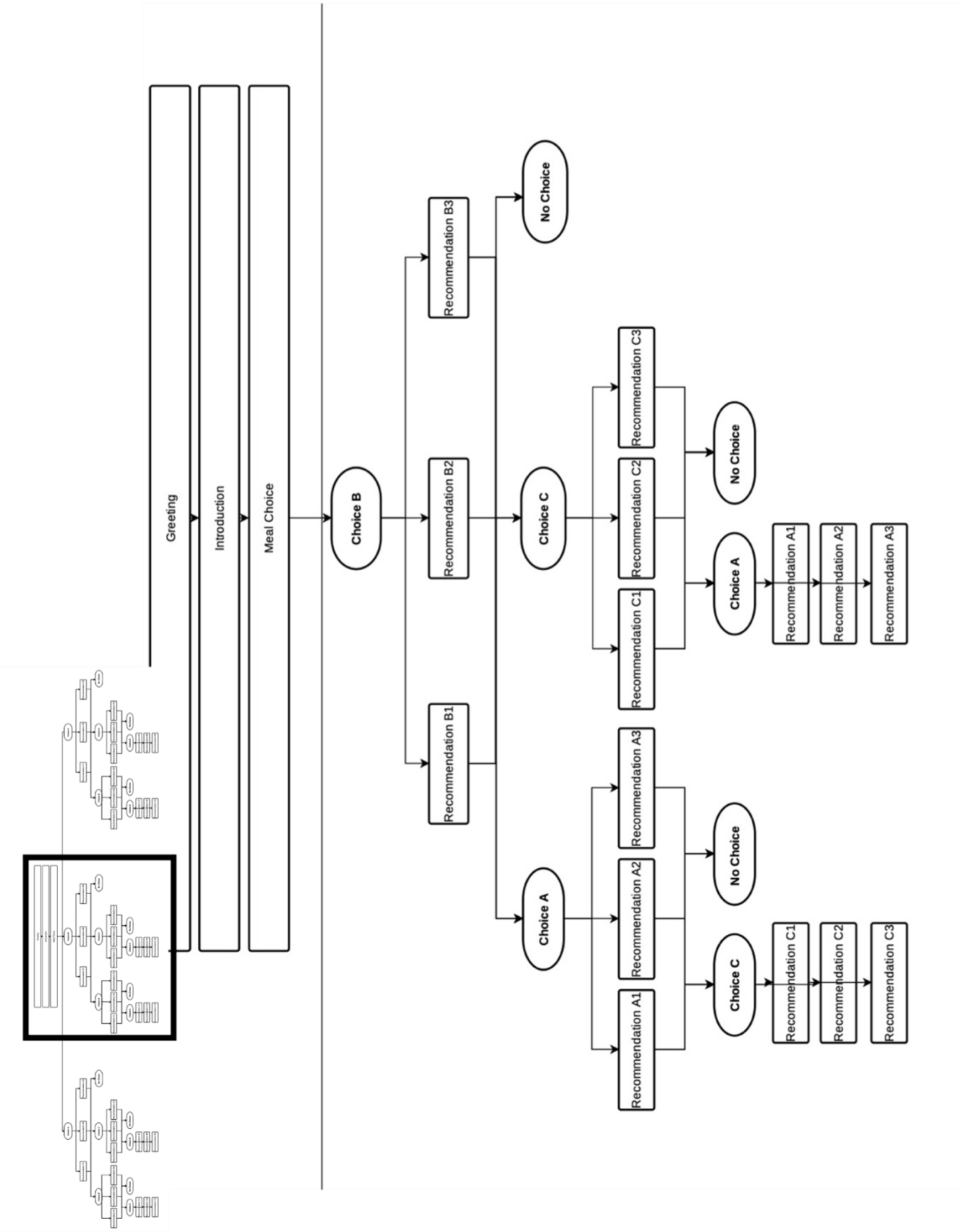
Zarantonello, L., & Pauwels-Delassus, V. (2017). *The Handbook of Brand Management Scales.* Routledge.

8 Appendix

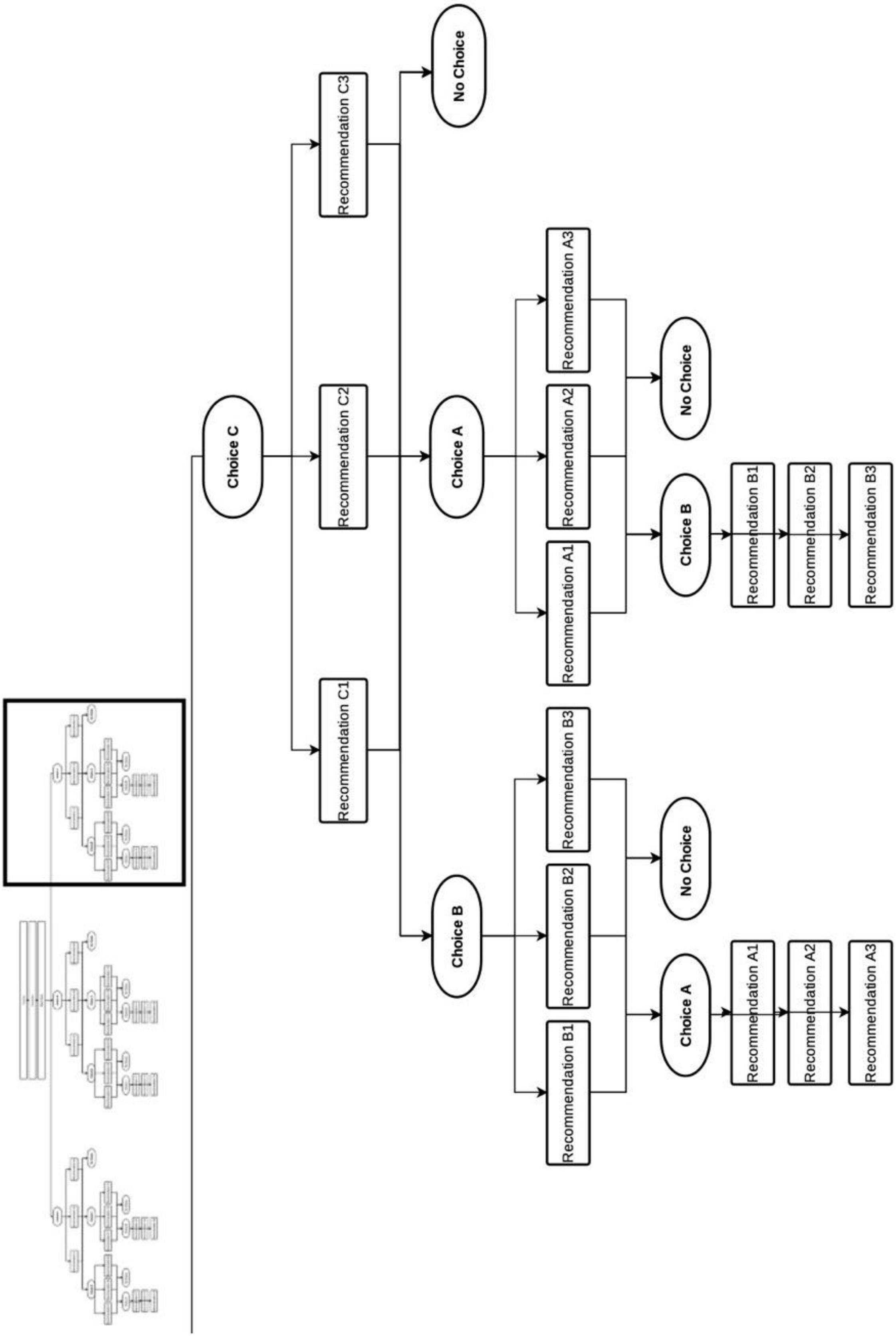
Appendix 1a): Chatbot System Design: Part A of the entire conversation path



Appendix 1b): Chatbot System Design: Part B of the entire conversation path



Appendix 1c): Chatbot System Design: Part C of the entire conversation path



Appendix 2: Questionnaire items

Variable	Item	No.	Ref.
Anthro-pomorphic design cues (ADC)	Using this chatbot gives me a feeling of personal communication.	ADC_1	Go & Sundar (2019) Appendix A
	Using this chatbot gives me a feeling of sociability.	ADC_2	
	Using this chatbot gives me a feeling of human warmth.	ADC_3	
	Using this chatbot gives me a feeling of human contact.	ADC_4	
	Using this chatbot gives me a feeling of human sensitivity.	ADC_5	
	Using this chatbot gives me a feeling of being present with someone.	ADC_6	
	This chatbot behaves human-like.	ADC_7	Nowak & Rauh (2005); Goetz et al. (2003)
	This chatbot shows human characteristics.	ADC_8	
	This chatbot behaves emotionally.	ADC_9	
Customer Satisfaction (CS)	I am satisfied with this chatbot.	CS_1	Chung et al.(2020)
	I am content with this chatbot.	CS_2	
	This chatbot did a good job.	CS_3	
	This chatbot did what I expected.	CS_4	
	I am happy interacting with this chatbot.	CS_5	
	I am satisfied with my overall experience using this chatbot.	CS_6	
	I am satisfied with the chatbot's service.	CS_7	
	I would recommend others to use this chatbot.	CS_8	
Perceived Enjoyment of the chatbot (PE)	Using this chatbot is fun.	PE_1	Mikalef et al.(2013)
	Using this chatbot is exciting.	PE_2	
	Using this chatbot is entertaining.	PE_3	
	Using this chatbot is delightful.	PE_4	
	Using this chatbot is enjoyable.	PE_5	
	Using this chatbot makes me happy.	PE_6	
Trust(-worthiness) towards (of) the chatbot (TW)	This chatbot cares about the consumer's needs.	TW_1	Zarantonello and Pauwels-Delassus (2016)
	This chatbot keeps its promises.	TW_2	
	This chatbot is able to contribute to the consumer's well-being.	TW_3	
	I have no doubt this chatbot can be trusted.	TW_4	
	This chatbot is trustworthy.	TW_5	
	I trust this chatbot.	TW_6	
Attitude towards the chatbot (AT)	This chatbot in the context of online shopping seems like a good idea.	AT_1	Venkatesh et al.(2003)
	This chatbot makes the webshop more interesting.	AT_2	
	Using this chatbot on the webshop is fun.	AT_3	
	This chatbot makes a friendly impression on me.	AT_4	Spears et al.(2004)
	Interacting with this chatbot is a positive experience.	AT_5	

Appendix 3: Demographic characteristics of respondents (after validation)

Variable	Category	N	%
Gender	Male	182	49.1%
	Female	189	50.9%
Age	14-24	120	32.3%
	25-40	251	67.7%
Education	Less than high school degree	4	1.1%
	High school degree or similar	70	18.9%
	Bachelor's degree	158	42.6%
	Master's degree / MBA or similar	136	36.7%
	Doctorate degree / PhD	3	0.8%
Job Level	Student	141	38.0%
	Working Student	45	12.1%
	Employed	173	46.6%
	Unemployed	12	3.2%
Nationality	German	292	78.7%
	Portuguese	19	5.1%
	Italian	22	5.9%
	Spanish	1	0.3%
	French	5	1.3%
	American	7	1.9%
	British	1	0.3%
	Dutch	2	0.5%
	Other	22	5.9%

Appendix 4: Online shopping characteristics (after validation)

Variable	Category	N	%
Previous Chatbot Interaction	Yes	319	86.0%
	No	28	7.5%
	I don't know	24	6.5%
Shopping Frequency	1-3 times a year	10	2.7%
	4-6 times a year	55	14.8%
	7-12 times a year	108	29.1%
	1+ per month	133	35.8%
	1-2 times per week	59	15.9%
	3+ per week	2	0.5%
	never	4	1.1%

Appendix 5: Results of the confirmatory factor analysis

Variable	KMO	Bartlett's Test	Result
ADC	.959	Approx. Chi-Square	5342.970
		df	36
		Sig.	.000
CS	.954	Approx. Chi-Square	4964.569
		df	28
		Sig.	.000
PE	.936	Approx. Chi-Square	3706.848
		df	15
		Sig.	.000
AT	.901	Approx. Chi-Square	3733.580
		df	15
		Sig.	.000
TW	.892	Approx. Chi-Square	2751.667
		df	10
		Sig.	.000

Appendix 6: Results of the reliability measurement

Chatbot	Scales	Cronbach's Alpha	Number of Items
Luigi	ADC	.884	9
	CS	.887	8
	PE	.887	6
	AT	.926	5
	TW	.910	6
Standard	ADC	.893	9
	CS	.850	8
	PE	.863	6
	AT	.870	5
	TW	.870	6

Appendix 7: Independent samples t-test results

Variable	Chatbot	N	M	SD	SEM
ADC	Luigi	186	5.36	1.049	.077
	Standard	185	2.68	1.122	.083
CS	Luigi	186	5.65	.889	.065
	Standard	185	2.96	1.261	.093
PE	Luigi	186	5.42	1.092	.080
	Standard	185	2.71	1.149	.085
AT	Luigi	186	5.47	1.148	.084
	Standard	185	3.02	1.398	.103
TW	Luigi	185	5.78	.799	.059
	Standard	185	2.80	1.272	.094

Appendix 8: Correlation tables

A) Chatbot Luigi

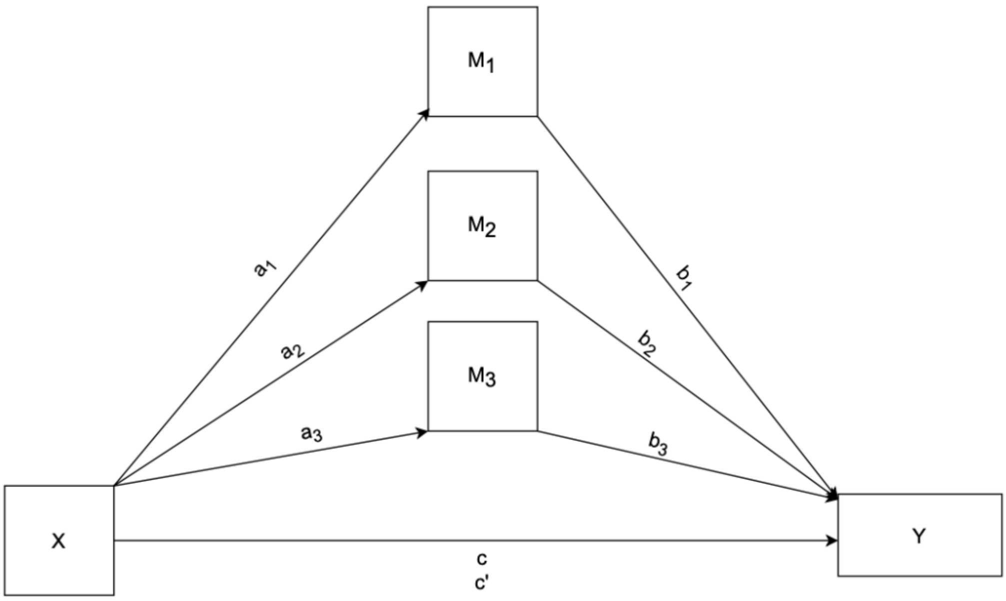
	Variable	N	M	SD	1	2	3	4	5
1.	Age	186	1.7	.48					
2.	ADC	186	5.36	1.05	-.175*				
3.	PE	186	5.42	1.09	-.195**	.800**			
4.	TW	186	5.47	1.15	-.204**	.728**	.688**		
5.	AT	185	5.78	.80	-.072	.625**	.672**	.457**	
6.	CS	186	5.65	.89	-.163*	.819**	.786**	.759**	.610**

B) Chatbot standard

	Variable	N	M	SD	1	2	3	4	5
1.	Age	185	1.7	.46					
2.	ADC	185	2.68	1.12	-.067				
3.	PE	185	2.71	1.15	-.041	.613**			
4.	TW	185	3.02	1.40	-.069	.505**	.626**		
5.	AT	185	2.80	1.27	-.027	.468**	.662**	.671**	
6.	CS	185	2.96	1.26	-.061	.616**	.711**	.719**	.687**

Note: **p< .01; *p< .05

Appendix 9: Applied mediation model



Appendix 10: Result output of mediator model in SPSS (PROCESS)

```
************ PROCESS Procedure for SPSS Version 3.5.3 ***********
Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
Documentation available in Hayes (2018). www.guilford.com/p/hayes3
******************************************************************
Model  : 4
Y  : Mean_CS_
X  : Mean_ADC
M1  : Mean_PE_
M2  : Mean_TW_
M3  : Mean_AT_
Sample
Size:  370
******************************************************************
OUTCOME VARIABLE:
Mean_PE_
Model Summary
R       R-sq        MSE      F(HC3)        df1        df2          p
.8800  .7744       .7020   1389.5613     1.0000   368.0000    .0000
Model
coeff       se(HC3)     t        p       LLCI       ULCI
constant    .4476    .1178   3.7998    .0002    .2160           .6792
Mean_ADC    .9002    .0241  37.2768    .0000    .8527           .9477
******************************************************************
OUTCOME VARIABLE:
Mean_TW_
Model Summary
R        R-sq      MSE      F(HC3)        df1        df2          p
.8153   .6647     1.0553     928.2954      1.0000   368.0000   .0000
Model
coeff        se(HC3)     t         p         LLCI       ULCI
constant   .8663    .1446   5.9897    .0000    .5819          1.1506
Mean_ADC   .8387    .0275  30.4679    .0000    .7846           .8928
******************************************************************
OUTCOME VARIABLE:
Mean_AT_
```

```
Model Summary
R         R-sq      MSE       F(HC3)         df1       df2           p
.8235   .6782      1.0775   1107.2803     1.0000   368.0000   .0000
Model
coeff        se(HC3)    t        p        LLCI     ULCI
constant   .7685      .1315   5.8440   .0000   .5099          1.0271
Mean_ADC   .8739      .0263   33.2758  .0000   .8223           .9256
*******************************************************************
OUTCOME VARIABLE:
Mean_CS_
Model Summary
R         R-sq      MSE    F(HC3)       df1       df2                p
.9382   .8802    .3643   815.6467   4.0000   365.0000          .0000
Model
coeff        se(HC3)    t          p        LLCI        ULCI
constant     .2280     .0923     2.4699   .0140       .0465      .4095
Mean_ADC     .2521     .0583     4.3213   .0000       .1374      .3669
Mean_PE_     .2321     .0596     3.8912   .0001       .1148      .3493
Mean_TW_     .2669     .0526     5.0715   .0000       .1634      .3704
Mean_AT_     .2294     .0470     4.8803   .0000       .1370      .3218
********************** TOTAL EFFECT MODEL ***********************
OUTCOME VARIABLE:
Mean_CS_
Model Summary
R        R-sq      MSE      F(HC3)         df1         df2           p
.8793  .7731    .6840    1366.9798     1.0000   368.0000       .0000
Model
coeff    se(HC3)      t         p           LLCI      ULCI
constant .7393      .1242    5.9514      .0000       .4950      .9836
Mean_ADC .8854      .0239   36.9727      .0000       .8383      .9325
********* TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y *********
Total effect of X on Y
Effect    se(HC3)      t       p       LLCI     ULCI     c'_ps       c'_cs
.8854   .0239    36.9727  .0000   .8383   .9325    .5106       .8793
Direct effect of X on Y
Effect    se(HC3)      t       p       LLCI     ULCI      c'_ps      c'_cs
```

```
.2521    .0583    4.3213   .0000    .1374    .3669    .1454     .2504
Indirect effect(s) of X on Y:
Effect     BootSE   BootLLCI   BootULCI
TOTAL         .6332      .0477      .5352      .7224
Mean_PE_      .2089      .0533      .1013      .3108
Mean_TW_      .2239      .0435      .1355      .3079
Mean_AT_      .2005      .0410      .1256      .2885
Partially standardized indirect effect(s) of X on Y:
Effect     BootSE   BootLLCI   BootULCI
TOTAL         .3652      .0268      .3102      .4162
Mean_PE_      .1205      .0307      .0584      .1805
Mean_TW_      .1291      .0250      .0783      .1774
Mean_AT_      .1156      .0235      .0726      .1652
Completely standardized indirect effect(s) of X on Y:
Effect     BootSE   BootLLCI   BootULCI
TOTAL         .6289      .0452      .5337      .7124
Mean_PE_      .2075      .0528      .1003      .3088
Mean_TW_      .2223      .0429      .1352      .3046
Mean_AT_      .1991      .0402      .1253      .2836
******* BOOTSTRAP RESULTS FOR REGRESSION MODEL PARAMETERS ********
OUTCOME VARIABLE:
Mean_PE_
Coeff   BootMean     BootSE   BootLLCI   BootULCI
constant      .4476      .4470      .1185      .2207      .6833
Mean_ADC      .9002      .9004      .0244      .8511      .9470
----------
OUTCOME VARIABLE:
Mean_TW_
Coeff   BootMean     BootSE   BootLLCI   BootULCI
constant      .8663      .8667      .1419      .5905     1.1439
Mean_ADC      .8387      .8386      .0269      .7855      .8922
----------
OUTCOME VARIABLE:
Mean_AT_
Coeff   BootMean     BootSE   BootLLCI   BootULCI
constant      .7685      .7667      .1323      .5072     1.0346
```

```
Mean_ADC      .8739      .8742      .0265      .8222      .9253
----------
OUTCOME VARIABLE:
Mean_CS_
Coeff   BootMean     BootSE   BootLLCI   BootULCI
constant      .2280      .2259      .0917      .0504      .4049
Mean_ADC      .2521      .2542      .0564      .1473      .3672
Mean_PE_      .2321      .2288      .0589      .1123      .3442
Mean_TW_      .2669      .2657      .0527      .1599      .3677
Mean_AT_      .2294      .2321      .0464      .1420      .3266
*********************** END MATRIX ************************
```

Declaration of Honor

I do solemnly declare that I prepared this thesis independently and that the thoughts taken directly or indirectly from other sources are indicated accordingly. This work has not been submitted to any other examination authority and further not yet been published.

Lisbon, September 07th, 2021

K. Klein

Katharina Klein