Anthropomorphic Service Recovery: the Panacea Following Service Failure of Automated Customer Service Agents (CSAs)

Nobukhosi Dlodlo*

Department of Marketing, Retail Business & Sport Management, Vanderbijlpark, South Africa, Orcid: 0000-0002-4727-5453

Keywords

Empathy

Humour

Service failure

Automated customer service agent

Abstract

This study leverages on the computer are social actors (CASA) theory to explore consumer response mechanisms to different types of service failures and recovery strategies of automated customer service agents. The paper asserts that selected anthropomorphic traits of conversational agents could make shoppers perceive them as having more human-like attributes such as humour and response empathy, thereby enhancing the trustworthiness of the non-human CSA. This study followed a scenario-based quantitative survey that was distributed as a survey hyperlink, comprising 287 participants. The findings showed a significant effect between empathy and anthropomorphism and perceived trustworthiness. In addition, anthropomorphic CSAs provide a trust shield effect, reducing the loss of trust following a service failure. Consequently, shoppers are more willing to forgive the online retailer. This study provides initial evidence that humans tend to be more forgiving of the failures of an anthropomorphic technological interaction partner. The findings of this study can enrich the response mechanisms and boundary conditions of online service failure by automated CSAs and provide important insights for online retailers to effectively respond to service failure and make reasonable use of human-robot collaborative work.

1. Introduction

1.1. Background

Thorbjørnsen *et al.* (2002:19) define interactive marketing as "an iterative dialogue where individual consumers' needs and desires are uncovered, modified and satisfied to the degree possible." In this respect, digital conversational assistants have tremendous potential to boost customer engagement, which is an essential metric in marketing as it may result in raised conversions and revenue (Kaczorowska-Spychalska, 2019). Chatbots for example, by their very conversational nature are continuously recognised as a viable marketing tool for establishing and fostering personal relationships with customers (Lekaviciute, Auruskeviciene & Reardon, 2023). While automated customer service agents excel in handling simple routine requests from customers (Chaves & Gerosa, 2018), they are somewhat limited when attempting to handle complex and nuanced scenarios, often yielding instances of online service failure.

1.2. Problem Statement

Endowing conversational CSAs with human-like characteristics improves communication and fosters social and emotional connections (Araujo, 2018; Chaves & Gerosa, 2018). Interaction with a human-like CSA eventually results in positive attitudes toward the conversational agent or the brand that it represents (Araujo, 2018). Be that as it may, conversational agents are not human, they are not even pretending to be human; save for that they only behave human-like rather than human (Lutz, 2023). However, the customers through interaction tend to perceive them as social actors, which somewhat yields unsettling sensations and higher expectations. In this respect, the literature is awash with studies that seek to focus on improving the design of automated CSAs to enhance their functional precision (e.g. Touré-Tillery & McGill, 2015). While this is so, Jain *et al.* (2023) as well as Đuka and Njeguš (2021) conceded in suggesting that conversational agents such chatbots should incorporate social capabilities. Further, Neururer *et al.* (2018) highlight that making a conversational agent likeable to users is essentially a social issue, not a technological one. The problem remains determining which social features are crucial for enhancing the communication and social abilities of conversational agents. This problem is prominent during periods when the acceptance of automated conversational agents by shoppers is low such as in case of service failure, where an even darker side of conversational agents is portrayed that could result in customer aggression (Huang & Dootson, 2022).

Scholars such as Frank and Otterbring (2023) and Adam *et al.* (2021) have identified the necessary qualities an AI-based front-line employee requires to service customers at the same level as human employees. Be that as it may, service failure research documents that the recovery strategies of automated customer service agents can affect people's psychological perceptions (e.g. Song *et al.* 2023). This research seeks *to advance the discourse in how automated customer service agents can seem more human, while demonstrating anthropomorphic characteristics, since such an understanding affects and is a necessary ingredient to deliver an alteration in shoppers' trustworthiness perceptions via forgiveness.*

1.3. Research objectives

This study comes in response to urges made by service scholars such as Song *et al.* (2023) to bridge the gap in the service failure-recovery literature by exploring the role of conversational agents in effective service recovery. The study had two objectives: First, to explore the anthropomorphised traits of automated CSAs (empathy and humour). Second, to explore the mediating role of trustworthiness and anthropomorphism on customer forgiveness, of which the latter is considered a calculated response to the marketer's service recovery efforts.

The paper is structured as follows: First, the background introduces the area of study, whereas the paper proceeds to outline the research gaps via problematisation and identification of clear lines of inquiry to pursue as the study objectives. Thereafter, a comprehensive review of the literature is done, to elucidate the suppositions made in this work. The CASA theory is explained, followed by the study hypotheses. This is followed by the research design and methodology, methods and sampling procedures, data analysis and presentation of the results. Finally, conclusions, limitations and avenues for future research are presented.

2. Literature Review

While they may be attributed with human-like names (Boldi *et al.* 2023) automated customer service agents are not infallible. Scenarios of interactive service failure might include the limited ability to understand the customer's voice tone, accent and other emotional nuances during communication. For instance, when a customer attempts to address a credit card billing query to finalise a purchase order, hurriedly they might encounter generic replies from the automated customer service agent, which might yield feelings of excessive irritation and discontent. Another problem is usually the absence of empathy and politeness in the reply. For example, upon attempting to report a possible fraudulent encounter on their mobile banking application, the programmed response from the automated customer service agent might seem to lack co-operation and sympathy. These failures underscore the constraints of using automated customer service agents, where the absence of personalisation and profound understanding of the customer's tone can deepen dissatisfaction.

2.1. Outcomes of service failure

Consumers respond negatively to service failure issues through the attributes of taking revenge and spreading nWOM (Casidy & Shin, 2015) and such responses adversely influence a firm's brand image. It is thus vital to understand these post-failure reactions of consumers in the technology context, where chatbots are responsible for service recovery. While it could be a significant gain for marketers if consumers forgive a brand for service failure and diminish their propensity to spread nWOM (negative word- of-mouth), research is silent on the chatbot traits that render them effective in managing online service failure recovery.

2.2. Computers as social actors (CASA) theory

The school of thought forward by Nass and Moon (2000:93) denotes Computers as Social Actors" (CASA), implying their anthropomorphistic nature. This implies that users are able to reflect on the human-like attributes of the automated agent during a social interaction, thereby inferring a more mindful psychological process (Boldi *et al.* 2023; Adolphs, 2010). It is this latter school of thought that is adopted in this research, further drawing from the central tenets of the theory of reasoned action.

2.3. The CASA theory and anthropomorphic traits of automated CSAs

If automated CSAs could appear like humans and reflect empathy in their text messages, they, as per CASA, are more likely to be perceived as social actors (Pelau *et al.,* 2021). Thus, anthropomorphic conversational agents depicting an empathetic and humorous communication style should help raise consumers' positive responses toward the company after a service failure. However, although chatbots could be perceived as social actors by their communication and appearance, consumers are also aware that chatbots are machines. Thus, their privacy concerns regarding personal information remain when interacting with chatbots. Overall, these three attributes of chatbots, i.e., empathy, anthropomorphism, and privacy concerns, influence consumers' perceived trustworthiness toward chatbots. Thus, perceived trustworthiness, where the consumer believes that a chatbot has securely taken service recovery steps after understanding customer problems, may calm the consumer's negative emotions.

2.3.1. Empathy

Customers are somewhat hesitant to interact with automated CSAs as the personal touch is lacking when the agent is non- human. Furthermore, as consumers incur loss of a varied nature such as economic, time and convenience losses during service failure, they also want an agent who could express empathy with the loss experienced by the customer, which is generally possible with humans (Nguyen *et al.,* 2022). For example, in a service failure context, if conversational agents could respond warmly and compassionately, it could elicit within consumers gloomy feelings caused by the service failure, thus making consumers perceive the CSAs to be benevolent and caring about customer issues (Rapp *et al.,* 2021). Additionally, the automated CSAs could also provide empathetic cues as they converse with consumers for service recovery, making consumers perceive them as empathetic.

H1: The positive effect of empathy on trustworthiness of automated CSAs is mediated by perceived anthropomorphism.

2.3.1.2. Humour

In human-human interactions, humour can be spontaneous yet instrumental in regulating a conversation and more so, in negotiating a service recovery process. Humour allows criticism to be smoothened and also acts as a stress-reliver. Humour can also help with frustration. According to Lekaviciute *et al.* (2023), in commercial settings, humour can be useful in inducing customer trust. By incorporating humour into chatbot interactions, businesses may be able to create a more engaging and memorable customer experience (Chaves & Gerosa, 2018). It is probable that humour, as an exceptionally human characteristic, increases anthropomorphism and, therefore, could theoretically increase connection and ultimate trustworthiness of the automated CSA. Thus:

H2: The positive effect of humour on trustworthiness of automated CSAs is mediated by perceived anthropomorphism.

2.3.1.3. Anthropomorphism

Compared to mindlessness which is considered an automatic, involuntary and spontaneous psychological process, anthropomorphism is perceived to involve "thoughtful, sincere beliefs that the object has human characteristics" (Nass & Moon, 2000, p. 93). Anthropomorphism refers to the process whereby individuals attribute humans' mental or emotional activities to non-human agents to interpret their actions (Epley et al., 2007 as cited in Van Pinxteren et al., 2020:206). In other words, anthropomorphism requires more reflective thinking, thereby inferring a more mindful psychological process (Adolphs, 2010). While research has attempted to examine whether it is mindlessness or anthropomorphism that drives users' social responses to technologies (Kim & Sundar, 2012; Lee & Oh, 2021), mixed findings have emerged, which steers the course of this paper.

2.4. Trustworthiness following an anthropomorphised service recovery effort

The most influential model of trust was provided by McKnight, Choudhury and Kacmar (2002), based on a three-dimensional conceptualisation of interpersonal trust. According to this conceptualisation, beliefs about the trustee's competence, integrity and benevolence determine the initial trust level. According to Lankton *et al.* (2015), for the trustworthiness of technology, the technology must fulfil three criteria; namely perceived ability (competence), benevolence and integrity. Perceived ability implies that with technology, a customer considers if it renders the assured performance (McKnight *et al*. 2002). Competence reflects a trustor's perception of a trustee's domain-specific knowledge and ability to successfully produce the desired outcome (Mayer, Davis & Schoorman, 1995). For example, the ability of an automated CSA to resolve an online shopping inquiry would be perceived as being able to deliver the promised. The benevolence dimension of trustworthiness reflects the extent to which the trustee is perceived to be motivated to always put the interests of the trustor first (Bhattacherjee, 2002). This aligns with individuals' hope that the automated CSA cares enough to offer help when needed. For technology, users also hope that a technology's help function will assist them with the information necessary to complete a task (McKnight *et al.* 2002). Integrity, finally, represents the ethical dimension of trustworthiness. It entails as the affinity to remain consistent in performance. This includes the persistent predictability of the automated CSA even during scenarios of service failure. By responding predictably to inputs (such as responding to queries or printing on command), technology influences the user's perceptions of technology integrity.

Of note as per the CASA theory, anthropomorphic appearance in the context of service failure may make consumers perceive automated CSAs as agents that are not cold and programmed but able to resolve problems (Polakow *et al.,* 2021). Thus, the anthropomorphic appearance will make shoppers perceive chatbots to be equally competent and trustworthy as humans (or better) in resolving service failures. Thus;

H3: An anthropomorphised service recovery effort by an automated CSA yields perceptions of trustworthiness by shoppers.

2.5. Forgiveness following an anthropomorphised service recovery effort

Once consumers encounter service failure, they experience emotional state changes. Several factors influence the way consumers deal with their disappointment. Scholars have reported that consumer personalities such as religiosity and spirituality influence their propensity to forgive the firm for service failure (Tsarenko & Tojib, 2011). Among firm-level efforts, extant literature has found that asking to be forgiven, giving voice to consumers, and offering compensation enhances consumers' propensity to forgive the firm (Harrison-Walker, 2019). Scholars have also found the role of perceived justice in consumers' willingness to forgive the firm after service failure including efforts by the organisation to offer both apologies and compensation (Casidy & Shin, 2015).

If anger and desire for revenge are initially low, due to a minor service failure, a weak relationship, or a positive recovery, forgiveness is arguably moot. Fetscherin and Sampedro (2019) define forgiveness as letting negative emotions waive off, resulting from the wrongdoing of oneself, others or situations. In this case, forgiveness helps individuals regain psychological balance and engage in constructive behaviour with the offender, in this case the retailer (Tsarenko & Tojib, 2011). Thus, forgiveness can provide a foundation for relationship restoration with the transgressing brand or company (Xie & Peng, 2009). For example, acts of apologising, accepting blame, providing a candid explanation and providing an alternative recovery action plan are all ways to encourage favourable interactional justice evaluations (Jung & Seock, 2017). Thus, managers should find ways to facilitate consumer forgiveness following service failures.

H4: An anthropomorphised service recovery effort by an automated CSA yields acts of forgiveness by shoppers.

2.6. The relationship between perceived trustworthiness and forgiveness

Customers' willingness to forgive a brand for service failure is critical. Extant literature also suggests that customers are sometimes willing to forgive a company for service failure due to perceived trustworthiness, even if the service recovery outcome is not up to the expectation. This willingness to forgive a service failure happens when customers are convinced that the service provider made efforts to resolve a service failure issue, i.e., they perceived that the agent was benevolent and integral in their effort to resolve the problem. In other words, perceived trustworthiness occurs when the shopper believes that the automated CSA has securely taken adequate service recovery steps after understanding the customer' s problems. Thus, a consumer might be more willing to pardon the organisation for service failure. Conversely, if an automated CSA is not able to resolve a service problem up to the expectation, trust is likely to be broken.

H5: The positive effect of anthropomorphism on forgiveness is mediated by shopper perceptions of trustworthiness of automated CSAs.

3. Research Methodology

This study applied a scenario-based approach which is popular among studies related to service failure and recovery (e.g. Park & Ha, 2016; Singh & Crisafulli, 2015). A scenario-based approach was deemed superior when compared to recall-based or retrospective-reporting approaches because of their inherent robustness and lack of sensitivity to memory lapses and rationalisation tendency. Furthermore, a scenario-based approach is less-time consuming and also less-prone to ethical and managerial challenges, when compared to the enactment of real-life setting service failure that is attributed to specific organisations. Initially, exploratory research with 47 advanced diploma students (19 male; 28 female) were conducted to test the content validity of the measuring instrument. This test also assisted in determining the sector where service failure by automated CSAs is prevalent, namely online retail shopping for fashion (hedonic) merchandise. An online survey was administrated using GoogleDocs. In the scenario, a shopper was experiencing a service failure as their ordered item [low-end sneakers] was not delivered by the promised date by Rai.com, a hypothetical e-retailer. Thabo, the chatbot of Rai.com, first tried to understand the failure issue that the customer was experiencing. Next, they apologised with the exclamation ‘Oopsie Daisy’ stating that there was an inventory miscount on the online store. Thabo apologised for the inconvenience, tracked a store with inventory and offered an alternate store for collection of the same item that had been purchased online, including free delivery with a free shoe polish gift. The customer agreed to an alternate delivery date and time.

3.1. Target population

The target population was restricted to participants between the ages of 18 and 60 years old, with access to the survey link using Whatsapp and other social media distribution pages. The researcher deemed a sample size of 300 adequate in this study due to time constraints and budgetary considerations, of which 287 were usable. A total of 13 submissions were discarded owing to straight lining effects. Furthermore, Malhotra *et al.* (2017) recommends that the size of the study sample should be at least 300 when conducting multivariate statistics. Purposive sampling were applied by targeting those individuals with previous experience and understanding of interacting with an automated CSA.

3.2. Measures

The study’s constructs were operationalised using multi-item scales adapted from previously published studies, as shown in Table 1. The measuring instrument was further modified to fit the study’s context.

Perceived empathy was measured using a five-item scale adopted from Croes and Antheunis (2021), whereas the three-item humour scale was adapted from Cline *et al.* (2003). These two constructs were anchored along Seven-point Likert scales ranging from “1” (strongly disagree) to “7” (strongly agree). On the other hand, the perceived anthropomorphism of automated CSAs were measured by asking participants to rate four adjectives: likable, sociable, friendly and personal, on a seven-point scale ranging from “1” (describes very poorly) to “7” (describes very well), following the guidelines by Araujo (2018) as well as Kim and Sundar (2012).

Perceived trustworthiness was measured using a six-item scale by (McKnight *et al.* 2002). The consumer forgiveness scale consists of two subscales, depicting the absence of negative responses (six items) and the presence of positive responses (six items). In the present study, both the sub-scales were found to be positively (0.442) and significantly related ($p<0.001$), which is consistent with extant literature on forgiveness (Harrison-Walker, 2019). Forgiveness literature also explicitly mentions that both the sub-components of forgiveness are "intertwined and therefore inseparable" (Harrison-Walker, 2019; p. 382), and researchers should conduct further analysis using the scale in its completeness and not as two sub-constructs. Seven-point Likert scales ranging from "1" (strongly disagree) to "7" (strongly agree) were used to measure the scale items.

3.3. Ethics

The measuring instrument was evaluated, whereas the data collection process was sanctioned by the university ethics committee.

3.4. Common methods bias

According to Podsakoff *et al.* (2003), common method bias is a critical issue in questionnaire-based single-survey studies. Following the procedures recommended by Podsakoff *et al.* (2003) and Lindell and Whitney (2001), both statistical and procedural remedies approaches were applied to mitigate common method bias. First, as a prevention measure, the scale items used in this research were adapted from pre-validated measures. Secondly, the survey was self-administered (online), whereas respondent anonymity was maintained and they also received assurance for the same. Next, effort was taken to administer the survey questions in a randomised fashion. Finally, a single-factor confirmatory factor analysis approach was conducted as a statistical measure. The CFA procedure revealed an extremely poor fit (Chi-square/df =13.26; RMSEA =0.418; CFI =0.577; TLI =0.512), indicating the least influence of common method bias.

4. Results and Findings

4.1. Sample description

The respondent profile analysis shows that there were more male participants (54.6%; n=157) than female participants (45.4%; n=130). The sample was largely youthful, with most of the participants (52.6%; n=151) indicating that they were between 20 and 39 years and also holders of a post-high school certificate (58%; n=166).

4.2. Model fit

The PLS-SEM algorithm was employed. Initially, the value of Standardized Root Mean Square Residual Value (SRMR) was assessed to establish model fit. Moreover, the d_G and d_ULS values were below the 95 percent bootstrapped quantile. Therefore, these values indicate a good model fit. In addition, the AVE values are higher than 0.50. As such, the model is considered acceptable in terms of reliability and validity.

4.3. Measure validation using confirmatory factor analysis

The first step in reflective measurement model assessment involves examining the indicator loadings. While outer loadings that are greater than 0.50 are indicative of significant observed variables along each latent variable and those loadings below 0.50 are candidates for deletion, Hair *et al.* (2018:675-676) denote that researchers uphold those loadings above 0.708 as they indicate that the construct explains more than 50 per cent of the indicator's variance, thus providing acceptable item reliability. The outer loadings in this study ranged between 0.674 and 0.903, which is above 0.50 at the minimum. Be that as it may the item "I enjoy a service recovery conversation that includes jokes from an automated CSA (0.699) and "I will not stop thinking about how I was wronged by the online retailer" (0.674) were slightly below 0.708 but a decision was taken to retain the items as it was observed that both items would not compromise the purity of the model since they did not cross-load (indicator loading highly on more than one factor). Moreover, the manifest variables loaded in the same direction hypothesised by the theory, signaling no need for deletion.

Table 1: Confirmatory factor analysis results

Descriptors		Descriptive statistics		Reliability statistics		Convergent validity statistics			
Construct	**Item wording**	**Mean (SD)**	**Skewness (Kurtosis)**	**Outer loadings**	**Cronbach's alpha coefficient (α)**	**Composite reliability (RHO**	**Dillion-Goldstein's (RHO a)**	**AVE**	**The square root of AVE**

Empathy (Croes and Antheunis 2021)	Said the right things to make me feel	4.813 (1.027)	0.899 (0.237)	0.823	0.683	0.769	0.789	0.567	0.753
	Responded appropriately to my			0.859					
	Came across empathetic			0.827					
	Said the right thing at the right time			0.713					
	Good at understanding my problem			0.789					
Humour Cline *et al.* (2003)	I prefer a service recovery situation where the automated CSA is free to express their sense of	3.496 (0.619)	-0.178 (-0.637)	0.773	0.715	0.718	0.875	0.597	0.773
	I enjoy a service recovery conversation that includes jokes from an automated CSA			0.699					
	I often read written jokes by automated CSAs			0.772					
Anthropomorphic (Araujo, 2018; Kim and Sundar, 2012)	This automated CSA is likeable	4.496 (0.919)	0.883 (0.637)	0.808	0.815	0.819	0.839	0.778	0.882
	This automated CSA is sociable			0.715					
	This automated CSA is friendly			0.813					
	This automated CSA is personal			0.864					
Trustworthiness (McKnight, Choudhury &Kacmar, 2002)	I believe that the agent would act in my best interest	4.072 (0.710)	0.599 (0.337)	0.753	0.803	0.822	0.843	0.717	0.847
	If I required help, the agent would do its best to help me			0.732					
	The agent is interested in my well-being, not just its own			0.825					
	The agent is truthful in its dealings with me			0.773					
	I would characterise the agent as honest			0.774					
	The agent would keep its commitments			0.870					
Shopper forgiveness	I will not stop thinking about how I was wronged by the eRetailer	4.229 (1.927)	1.043 (0.106)	0.674	0.901	0.901	0.937	0.788	0.887

(Harrison-Walker, 2019)	I will spend time thinking about how to get back at the eRetailer for the service failure			0.801					
	I will avoid certain websites because they will remind me of the e-retailer who wronged me. (R)			0.804					
	This e-retailer's wrongful actions will keep me from enjoying life.			0.715					
	(R) I think that many of the emotional wounds related to the e-retailer's wrongful actions will heal.			0.829					
	I think my life will be ruined because of the e-retailer's wrongful actions. (R)			0.740					
	I wish for good things to happen to the e-retailer who wronged			0.816					
	If I encounter the e-retailer who wronged me, I will feel at peace.			0.824					
	I have compassion for the e- retailer who wronged me.			0.700					
	I hope the e-retailer who wronged me is treated fairly by others in the future.			0.691					
	I forgive the e-retailer for what they did to me.			0.903					
	Even though the e-retailer's actions hurt me, I have goodwill for the e-retailer.			0.755					
Thresholds				0.708	0.70	0.70	0.70	0.50	Higher than the correlation coefficients

Source: Author's compilation

4.4. Reliability assessment

Cronbach alpha values shown in Table 1 authenticate that each construct accomplished the minimum cut-off value of 0.70 (Malhotra *et al.* 2017), save for empathy which reported a Cronbach's alpha coefficient of 0.683. Be that as it may, Sarstedt *et al.* (2021:17) acknowledge that while Cronbach's alpha coefficient assumes the same thresholds composite reliability, it generally yields lower values. The study proceeded noting this together with the submission by Cohen *et al.* (2018:774), who pointed out that values between 0.60 and 0.69

can be interpreted as evidence of marginal reliability. Other statistics that were computed and proved the study to be reliable include the composite reliability statistic (Ranging between 0.718 and 0.901) as well as Dillion-Goldstein' rho statistic (Ranging between 0.789 to 0.937), which were above 0.70.

4.5. Convergent and discriminant validity

The average variance extracted (AVE) values are higher than 0.50, depicting convergent validity. As such, the model is considered acceptable in terms of reliability and validity. Table 2 reports on the convergent and discriminant validity assessment for this research.

Table 2: Discriminant validity results

Latent variable	Empathy	Humour	Anthro	Trustworthiness	Forgiveness	HTMT
Empathy	**0.753**					0.237
Humour	0.248**	**0.773**				0.841
Anthro	0.329**	0.489**	**0.882**			0.134
Trustworthiness	0.235**	0.686**	0.387**	**0.847**		0.577
Forgiveness	0.335*	0.316**	0.245**	0.510**	**0.887**	0.658

Note: Square roots (AVEs) are on diagonal, and construct correlations are below the diagonal AVEs of formative indicators are not applicable.

Source: Author's compilation

Three methods were employed to assess the discriminant validity, including evaluating the correlation coefficients, applying the Fornell-Larcker heuristic as well as the hetrotrait-monotrait correlation estimation. The highest correlation coefficient in the Pearson correlation matrix is 0.686 ($p < 0.01$), which is below 0.70 thereby depicting the theoretic uniqueness and seperability of each construct in the study. The highest correlation coefficient is lower than all the square root values of the AVEs associated with each construct. Furthermore, the HTMT values range between 0.139 and 0.841, which is below the 0,85 cut-off (Henseler *et al.* 2015:9-18). These three measures fulfilled the recommended criteria and established a satisfactory discriminant validity for all constructs in this study.

5. Structural model assessment

Table 3 depicts the results of the structural model assessment.

Table 3: Structural model results and hypotheses testing

Path			Hypotheses	Path coefficients (*β*)	T- statistic	Significance level	Decision
Perceived empathy	←	Anthropomorphism	H_1	0.339 (+)	2.890	***	H_{1a} is supported
Humour	←	Anthropomorphism	H_2	0.152 (+)	4.467	*	H_{1b} is partially supported
Anthropomorphism	←	Trustworthiness	H_3	0.438 (+)	5.859	***	H_{1c} is supported
Anthropomorphism	←	Shoppers' forgiveness	H_4	0.227 (+)	3.947	***	H_{1d} is supported
Perceived trustworthiness	←	Shoppers' forgiveness	H_5	0.269 (+)	3.805	***	H_{1e} is supported
Thresholds				**≥+0.20**	**>2.58**	**p<0.01**	**Hypotheses are supported by the sample data**

Source: Author's compilation

As per Table 3, all the latent variables had VIF values below 5.0, whereas the Tolerance values ranged between 0.298 and 0.678, which is above 0.10; thus dismissing any multicollinearity risk. Furthermore, inter-factor correlations, as reported in Table 2 were below 0.80 and hence, do not signify collinearity problems (Henseler *et al.* 2015). Moreover, a consistent bootstrapping procedure with 5000 subsamples was employed to calculate the path estimates and t-values. Table 3 provides detailed results from hypothesis testing. All the relationships are significant except for the impact of humour on shoppers' perceptions of anthropomorphism. In the study, at least 59.6 percent of the variance in anthropomorphism is explained by humour and empathy, whereas 54.1 percent of the variance in forgiveness is explained by both anthropomorphism and trustworthiness perceptions.

6. Theoretical and managerial Implications

This paper is a first step towards re-considering automated CSAs as social actors in a service failure scenario by acknowledging the transferability of human principles and/or the essence of anthropomorphism (demonstrated by incidents of chatbot empathy and humour) in triggering high perceptions of trustworthiness and ultimately, forgiveness inclinations among online shoppers. As a theoretical contribution, the findings of this study are crucial in as far as they exposed gaps in the service recovery strategies that relate to technology-based front-line employees. A This study provides fruitful opportunity to consider more developed research in the area of neuroscience and how conversational agents can be harnessed to alter consumer perceptions and emotions during shopping encounters. For managers, this study proved that it is vital for information systems and marketing managers to explore the traits of automated CSAs that could diminish a customer's negative response in case of service failure.

7. Conclusions, Limitations and Future Research

While scholars such as Ge and Gretzel (2017) found positive effects of humour on customer engagement in social media, this has been difficult to prove in a more consequential task-based scenario such as in online shopping, hence the diminishing indirect effect between humour and anthropomorphism. While it seems natural to assume that if humour improves human-human connection, comparable impacts can be expected on interactions between humans and conversational agents, this study could only partially prove the clinical validity of this indirect effect between humour and shoppers' perceptions of the trustworthiness of automated CSAs. The reason could be because the economic and non-economic losses suffered by customers when they engage in online shopping and fail to complete a shopping task seems to far outweigh any form of probable reparation that could be delivered through humour and jokes.

The scope of this study was restricted to a single-cross sectional survey hence the generalisability of the findings can be challenged. This study could be expanded to a larger sample and cross-country locations in view of enhancing the cultural validity of the findings. Again, the study adopted a mono-quantitative research methodology, which left no opportunity for triangulation of the study findings. Despite these shortcomings, this study may be used as the springboard by future researchers who may want to explore further the nexus between anthropomorphic traits in various service failure scenarios.

References

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, 31(2), 427-445.

Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences*, 1191, 42-61.
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.
Bhattacherjee, A. (2002). Individual trust in online firms: Scale development and initial test. *Journal of management information systems*, 19(1), 211-241.
Boldi, A., Curti, L., Perrucci, A., & Simeoni, R. (2023). How do people ascribe humanness to chatbots? An analysis of real-world human-agent interactions and a theoretical model of humanness. *International Journal of Human-Computer Interaction*, 1-24.
Casidy, R. & Shin, H. (2015). The effects of harm directions and service recovery strategies on customer forgiveness and negative word-of-mouth intentions. *Journal of Retailing and Consumer Services*, 27, 103-112.
Chaves, A. P., & Gerosa, M. A. (2018). How should my chatbot interact? A survey on human-chatbot interaction design. Pre-print arXiv, 190402743.
Cline, T. W., Altsech, M. B., & Kellaris, J. J. (2003). When does humor enhance or inhibit ad responses? – The moderating role of the need for humor. *Journal of Advertising*, 32(3), 31-45.
Cohen, L., Manion, L. & Morrison, K. (2018). *Research Methods in Education*. 8th ed. Oxon: Routledge.
Croes, E. A., Antheunis, M. L., van der Lee, C., & de Wit, J. M. (2024). Digital confessions: The willingness to disclose intimate information to a chatbot and its impact on emotional well-being. *Interacting with Computers*, iwae016.
Đuka, I., & Njeguš, A. (2021). Conversational Survey Chatbot: User Experience and Perception. In Proceedings of the International Scientific Conference Sinteza 2021.
Epley, N., A. Waytz, & J. T., Cacioppo. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4): 864-886.
Fetscherin, M., & Sampedro, A. (2019). Brand forgiveness. *Journal of Product & Brand Management*, *28*(5), 633-652.
Frank, D. A., & Otterbring, T. (2023). Being seen… by human or machine? Acknowledgment effects on customer responses differ between human and robotic service workers. *Technological Forecasting and Social Change*, *189*, 122345.
Ge, J., & Gretzel, U. (2017). The role of humour in driving customer engagement. In *Information and Communication Technologies in Tourism 2017: Proceedings of the*

International Conference in Rome, Italy, January 24-26, 2017 (pp. 461-474). Springer International Publishing.

Hair, J.F., Risher, J.J., Sarstedt, M., & Ringle, C.M. (2018). When to use and how to report the results of PLS-SEM*, European business review*, 31(1), 2-24.

Harrison-Walker, L. J. (2019). The critical role of customer forgiveness in successful service recovery. *Journal of Business Research*, 95, 376-391.

Henseler, J., Ringle, C.M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modelling. *Journal of the Academy of Marketing Science,* 43(1), 115-135.

Huang, Y. S. S., & Dootson, P. (2022). Chatbots and service failure: When does it lead to customer aggression. *Journal of Retailing and Consumer Services*, 68, 103044.

Jain, N. R. K., Liu-Lastres, B., & Wen, H. (2023). Does robotic service improve restaurant consumer experiences? An application of the value-co-creation framework. *Journal of Foodservice Business Research*, 26(1), 78-96.

Jung, N. Y., & Seock, Y. K. (2017). Effect of service recovery on customers' perceived justice, satisfaction and word-of-mouth intentions on online shopping websites. *Journal of Retailing and Consumer Services*, 37, 23-30.

Kaczorowska-Spychalska, D. (2019). How chatbots influence marketing. *Management*, 23(1), 251-270.

Kim, Y., & Sundar, S. S. (2012). Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior*, 28(1), 241-250.

Lankton, N. K., McKnight, D. H., & Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*, *16*(10), 1.

Lee, A. & Oh, H. (2021). Anthropomorphism and its implications for advertising hotel brands. *Journal of Business Research*, 129, 455-464.

Lekaviciute, V., Auruškevičienė, V., & Reardon, J. (2023). Should Your Chatbot Joke? Driving Conversion Through the Humour of a Chatbot Greeting. *Organizations and Markets in Emerging Economies*, 14(2 (28)), 220-241.

Lindell, M. K. & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs, *Journal of Applied Psychology*, 86(1), 114-121.

Lutz, C. (2023). Privacy and human-machine communication. *The sage handbook of human-machine communication*, 310-317.

Malhotra, N. K., Nunan, D., & Birks, D. F. (2020). *Marketing research*. Pearson UK.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review*, 20(3), 709-734.

McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). The impact of initial consumer trust on intentions to transact with a web site: a trust building model. *The journal of strategic information systems*, *11*(3-4), 297-323.

Nass, C., & Moon, Y., (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81-103.

Neururer, M., Schlögl, S., Brinkschulte, L., & Groth, A. (2018). Perceptions on authenticity in chat bots. *Multimodal Technologies and Interaction*, 2(3), 60.

Park, J., & Ha, S. (2016). Co-creation of service recovery: Utilitarian and hedonic value and post-recovery responses. *Journal of Retailing and Consumer Services*, 28, 310-316.

Pelau, C., Dabija, D. C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, 122, 106855.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5), 879.

Polakow, T., Laban, G., Teodorescu, A., Busemeyer, J. R. & Gordon, G. (2022). Social robot advisors: effects of robot judgmental fallacies and context. *Intelligent Service Robotics*, 15(5), 593-609.

Sarstedt, M., Ringle, C.M., & Hair, J.F. (). Partial least squares structural equation modeling. *In* Homburg, C., Klarmann, M. & Vomberg, A. *Handbook of market research.* Cham: Springer International Publishing. 587-632.

Singh, J., & Crisafulli, B. (2015). Customer responses to service failure and recovery experiences. *Boundary Spanning Elements and the Marketing Function in Organizations: Concepts and Empirical Studies*, 117-135.

Song, M., Xing, X., Duan, Y., & Mou, J. (2023). I can feel AI failure: the impact of service failure type and failure assessment on customer recovery expectation. *Industrial Management & Data Systems*, 123(12), 2949-2975.

Thorbjørnsen, H., Supphellen, M., Nysveen, H., & Egil, P. (2002). Building brand relationships online: A comparison of two interactive applications. *Journal of interactive marketing*, 16(3), 17-34.

Touré-Tillery, M., & McGill, A. L. (2015). Who or what to believe: Trust and the differential persuasiveness of human and anthropomorphized messengers. *Journal of Marketing*, 79(4), 94-110.

Tsarenko, Y., & Rooslani Tojib, D. (2011). A transactional model of forgiveness in the service failure context: a customer-driven approach. *Journal of Services Marketing*, *25*(5), 381-392.

Xie, Y., & Peng, S. (2009). How to repair customer trust after negative publicity: The roles of competence, integrity, benevolence, and forgiveness. *Psychology & Marketing*, *26*(7), 572-589.