

Research Article

Multi-Chatbot or Single-Chatbot? The Effects of M-Commerce Chatbot Interface on Source Credibility, Social Presence, Trust, and Purchase Intention

Su-Mae Tan ¹ and Tze Wei Liew ²

¹Faculty of Information Science and Technology, Multimedia University, Jalan Ayer Keroh Lama, 75450 Melaka, Malaysia

²Human-Centric Technology Interaction SIG, Faculty of Business, Multimedia University, Jalan Ayer Keroh, Lama, 75450 Melaka, Malaysia

Correspondence should be addressed to Tze Wei Liew; twliew@mmu.edu.my

Received 11 October 2021; Accepted 13 December 2021; Published 31 January 2022

Academic Editor: Zheng Yan

Copyright © 2022 Su-Mae Tan and Tze Wei Liew. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most chatbot interfaces in contemporary m-commerce platforms feature a single chatbot that provides recommendations for all product categories. Nonetheless, there is an emerging research interest in multi-chatbot systems designating multiple chatbots as product/domain-specific advisers. Given the dearth of studies investigating the effects of multi-chatbot versus single-chatbot in the m-commerce context, we addressed this research gap by conducting an online between-subjects experiment to explore how the m-commerce chatbot interface types can differently influence source credibility, social presence, trusting beliefs, and purchase intention. Based on 154 valid responses, the single-chatbot interface led to higher social presence and trusting beliefs toward the m-commerce platform than the multi-chatbot interface. Males attributed the chatbot with higher competence and reported higher purchase intention through the m-commerce platform when engaging with the single-chatbot interface than the multi-chatbot interface. These findings suggest that designating chatbots as product-specific advisers in a multi-chatbot interface without labels to accentuate expertise could not evoke the users to categorize them as product specialists. Moreover, the multi-chatbot interface could have imposed user confusion and unfamiliarity cues, decreasing trust in the m-commerce platform. These findings' theoretical, design, and managerial implications are discussed through the lens of the computers-are-social-actors paradigm, source credibility theory, source specialization, multiple source effect, and m-commerce behavioral research.

1. Introduction

Chatbots are computer programs that utilize text-based dialogue systems to simulate conversational engagement with humans [1–3]. Today, chatbots are popularly used in mobile platforms, fueled by the advances in artificial intelligence (AI) and natural language processing technology [4] and the significant business trend toward conversational commerce [5–8]. Chatbots have been deployed to perform various functions, including taking product orders from customers, providing answers to frequently asked questions, and dispensing product recommendations across m-commerce domains such as food (Pizza Hut), apparel (Zalora), travels (Malaysia Air-

lines), banking services (Bank of America), and multi-categorical products (eBay). By integrating chatbots within readily available chat interfaces familiar to users such as Facebook Messenger, Slack, WhatsApp, and WeChat [3], companies can provide automated personas that are ever ready to receive queries and dispense information anywhere anytime [8]. Currently, some of the popular product recommender chatbot agents implemented in mobile messaging applications are eBay ShopBot, Chatbot Sephora, Chatbot Castorama, and Chatbot H&M. These chatbots simulate the role of a product advisor by first asking a series of questions to users about their preferences, followed by offering personalized product recommendations based on users' responses.

Despite the increasing popularity of chatbot technology in m-commerce platforms, reports indicated that chatbots were not always well received by users. For instance, 80% of consumers believed that the after-sales services provided by chatbots were inferior to sales services offered by humans [9]. The same report also revealed that while 90% of the consumers thought businesses were ready to implement chatbots, 54% indicated that they would prefer to talk to human agents rather than chatbots. This issue highlights the importance of social and behavioral aspects of chatbot design from human-computer interaction (HCI) perspective [1, 10, 11]. As intention to accept and use a chatbot's recommendations depends on the user's social perceptions of the chatbots' credibility, social presence, and trust in the online platforms, it is thus crucial for chatbots to be designed to accentuate the qualities above [1, 2, 6, 11]. Drawing on the the computers-are-social-actors paradigm which states that human-to-human social rules apply to interaction between users and technological artifacts, chatbot designs should be considered carefully to create social cues evoking appropriate users' social and behavioral reactions during the engagement.

In general, contemporary multi-product category m-commerce platforms feature a single chatbot persona that provides recommendations for all product categories, such as eBay's Shopbot, Aerie, Spring Bot, and H&M bot. However, recent studies have considered the development of multi-agents and multi-chatbot platforms. For instance, three embodied virtual agents were dedicated to advising on each of the three product category types in multi-product categorical e-commerce websites [12, 13]. Researchers [14] devised three unique chatbots to advise on specific investment fields —PoupancaGuru as savings expert, TDGuru as treasury bonds expert, and CBDGuru as government bonds expert within a financial advisory platform known as Finch. Accordingly, users can direct their questions to the different chatbots to match the financial investment domain queries with the relevant domain-specific bots. A study [15] operationalized domain-specific chatbots conveying tourism information. The chatbots were assigned to specific travel domains, including nature (the coast, beaches, national parks, etc.); culture (architecture, museums, theaters, and history); and shopping (handmade crafts, malls, and souvenirs). Another research [16] reported the proposed implementation of six distinct embodied virtual agents in a virtual coaching system in the Council of Coaches project. The six embodied virtual agents would take on the specific role of an expert coach for each of the respective health domains: social, cognitive, physical, mental, diabetes, and chronic pain.

Notwithstanding the increased attention toward multi-chatbot systems in the emergent literature, most studies have focused on technical and system factors for establishing multi-chatbot platforms while under-exploring the social, affective, and behavioral effects of interacting with multi-chatbot systems [17–21]. This study, therefore, extends prior works on the multi-chatbot interface to *m-commerce*, which is one of the domains that utilize chatbot technology extensively. To the best of our knowledge, no studies have yet to

examine the social and behavioral effects of a multi-chatbot interface designating chatbots as product-specific advisors in an m-commerce platform. Hence, this study aims to fill this research gap by addressing the following questions.

RQ₁: To what extent do a multi-chatbot and single-chatbot m-commerce interface differently affect perceived chatbot competence[a], trust toward the m-commerce platform (ability, benevolence, and integrity) [b], intention to purchase through the m-commerce platform [c], and perceived social presence [d]?

RQ₂: To what extent does a user's gender moderate the effects of m-commerce chatbot interface (multi-chatbot and single-chatbot) on perceived chatbot competence [a], trust toward the m-commerce platform (ability, benevolence, and integrity) [b], intention to purchase through the m-commerce platform [c], and perceived social presence [d]?

2. Literature Review

2.1. Theoretical Perspectives on Designating Chatbot as Product-Specific Advisers in a Multi-Chatbot Interface. According to the computers-are-social-actors paradigm, people treat technological artifacts using the same social rules derived from human-to-human interaction [22]. It has been postulated that people's engagements with technological entities are mindless and automatic, as they rely upon overlearned social scripts sourced from their routine social experiences with people [23]. On this basis, technological entities can be devised to hold social cues activating a wide array of learned social scripts in the mind of users engaging with these entities, which influences users' perceptions and behavioral intentions toward the systems [24]. Among the different social cues, this study focuses on *specialty cues* [25, 26].

Specialty cues underpin this research as framing chatbots as product/domain-specific advisers can convey social cues which evoke users to categorize the agents as product or domain specialists. Specialty cues can elicit a heuristical assumption that social objects performing tasks in niche domains are specialists and, therefore, hold higher domain expertise and knowledge than generalists performing tasks across diverse domains [27–29]. For instance, the social descriptors denoting a “brain surgeon” and a “professor of Korean history” would cause people to perceive these social models as specialists within their niche domains [27]. This study draws on the *single functionality as a specialty cue hypothesis*, which posits that assigning information sources to single functionality can act as a cue to trigger specialization schema [28].

A study [28] tested the single functionality as a specialty cue hypothesis by posting notes on three IoT devices to designate each to uniquely convey weather, traffic, or event information, respectively. The IoT devices assigned to unique functions evoked stronger social presence, higher perceived expertise, and more positive attitudes toward the devices than the IoT devices that shared the same function of dispensing information on weather, traffic, or event. Other studies found similar effects with embodied virtual agents in multi-product category websites; specifically, users

ascribed higher perceived expertise toward the agents assigned to specific product categories than the single generalist agent dispensing recommendations for all product categories [12, 13]. However, it is worth noting that the researchers infused the product-specific agents with labels accentuating product specialization, for instance, “Hi, I am Anna — your camera product specialist,” which might have helped prompt the users to categorize the agents as product specialists.

Conversely, the single functionality as a specialty cue effect was not found in other studies. In an automated financial advisory system known as Finch, researchers assigned three unique chatbots as domain-specific experts in financial investment, including PoupancaGuru as savings expert, TDGuru as treasury bonds expert, and CBDGuru as government bonds expert [14]. In contrast, the generalist chatbot, InvestmentGuru, was tasked to advise all investment domains. However, the users did not rate the m-commerce platform’s competence, trust, and effectiveness differently between the domain-specific chatbot and single-chatbot design interfaces.

Through a Wizard of Oz experiment, researchers [15] found no differences in terms of conversation content, user’s speech, and impression toward the system when comparing between the tourism-related multi-bot system that featured domain-specific chatbots (Nature Agent, Culture Agent, and Shopping Agent) and the single-bot platform that had only one generalist agent (Natal Agent). Additionally, users reported more confusion in multi-chatbot engagements and utilized specific strategies to organize turn-taking when conversing with the chatbots.

The confusion within the user-agent interaction corroborates the argument made in the study [30] that users potentially prefer to engage with a consistent source rather than many diverse sources, as the former produces less cognitive load than the latter. Pertinently, the study did not find any main social effects between the multi-agent interface featuring different voice agents for each platform (smartphone, personal cloud computing, smart TV) and the single-agent interface incorporating one voice agent embodying these devices. Nonetheless, the user’s gender moderated the effects — females preferred the single consistent agent, whereas males preferred diverse voice agents.

Researchers [31] endowed one unique voice agency for each of the three different devices (TV, speaker, lamp) and found that the multi-agent interface did not produce main effects for social attraction and trust than the single-agent interface with one voice embodying all the devices. However, the user’s personality was a moderating factor insofar as extroverts reported higher social attraction and trust toward the media technology when interacting with a single-agent over the multi-agent frame. Whereas, introverts had stronger social attraction and trust toward the media technology with multiple-agents over the single-agent frame. The researchers argued that extroverts preferred interacting with the “more talkative” agent affording a longer conversation duration from the single voice embodying all the devices compared with the multi-agent interface. Introverts might have preferred interacting with the multi-agent interface

because interactions with each agent were briefer due to the dispersed agencies across the different devices compared to the single-agent interface.

2.2. Source Credibility Theory and Trust toward M-Commerce Recommender Platform. The source credibility theory attempts to explain how dimensions of an information source can influence users’ acceptance and use of the message conveyed by the source [32]. The dimensions of source credibility can be distinguished as trustworthiness and expertise [33, 34]. Trustworthiness refers to the extent to which the message source is perceived to be objective and honest; whereas, expertise refers to the degree to which the message source is perceived to hold high competence, skills, and knowledge required to dispense quality information [32, 35, 36]. This study draws on the literature surrounding source credibility and the social design of recommender systems, given that the chatbots are product recommenders in this study. Researchers [37] have asserted that technological sources conveying recommendations (recommender platforms) are social actors; therefore, users judge the source credibility of the systems through simple inspections based on exhibited social cues. As people are more likely to accept recommendations from credible sources, the credibility of recommender platforms is crucial in increasing the likelihood of message acceptance [38, 39]. Researchers [37] distinguished the credibility dimensions of recommender platforms to *expertise* and *trustworthiness*. Perceived *expertise* of a technological source refers to the extent to which the recommender system is perceived to have the ability, competence, and domain knowledge to perform its task effectively [40, 41]. On the other hand, the *trustworthiness* of a technological source is defined as 1) fair and unbiased and 2) motivated to prioritize the user’s interest over own gain.

Scholars [41] have posited that the source credibility or reputation of a website/vendor could consequently influence the user’s trust toward the recommender agent through a process known as trust transference. That is, buyers rely on cues associated with trusted “proof source” when dealing with unfamiliar sellers [42]. The transference process can also manifest in reverse insofar as the source credibility of agents representing the website/vendor (chatbot or virtual agent) can affect users’ trusting beliefs toward the platform that affords the agents [11, 13, 43]. Thus, the source credibility of agents translates into trusting beliefs toward the recommendation platform — trust ability, trust benevolence, and trust integrity [41, 44–46]. Trust ability refers to the user’s willingness to accept recommendations based on the perceived competence or expertise of the recommender platform. Trust benevolence is associated with the user’s perception that the recommender system cares and acts in the user’s interest rather than prioritizing profit on behalf of vendors. In contrast, trust integrity refers to the extent to which a user perceives the recommender system as truthful and ethical in providing information and fulfilling its promises [41]. Given that recommender systems are treated as social actors, such conceptualization of trust factors is crucial insofar that even when users find these sources to be

competent, they may not fully trust these sources for they may believe that these sources hold their motivation and goals, which are prioritized over user's benefits [14, 45].

2.3. Purchase Intention through the M-Commerce Platform. Social cues of anthropomorphic agents in websites can drive users to purchase through the platform [11, 43, 46]. Relatedly, the source credibility of product speakers and celebrity endorsers positively affects purchase intention in digital commerce [36, 47–49]. The effects of source expertise on purchase intention can occur with technological sources conveying social cues, as informed by the computers-are-social-actors paradigm. A recent study has shown that chatbots' perceived credibility and competence contributed to higher purchase intention [11]. Another study [43] revealed that embodied virtual agents resembling product advisors whose gender (male/female) matched with their product gender type (feminine/masculine) elevated perceived expertise of technological source and purchase intention through the website. Moreover, an experiment [27] attached labels to infuse specialty cues into smartphone hardware and application, and the cues enhanced the user's purchase intention toward the advertised product. In the context of a multi-product category e-commerce website, users reported higher purchase intention through the web stores featuring virtual agents framed as product-specific advisors [12, 13].

The literature states that trusting beliefs toward the platform consequently drive users' intention to purchase through the platform [50, 51]. In a study examining the social presence effects of an embodied virtual agent in a website [46], only trust ability of the website (and not trust benevolence and trust integrity) led to higher patronage intention. An experiment comparing a multi-agent versus a single-agent interface demonstrated that website trust ability and trust benevolence were significant predictors of intention to purchase through the website [12].

2.4. Social Presence and Multiple Source Effect. The role of social presence is crucial in digital commerce, as it drives positive reactions of users toward website trust and purchase/patronage intention [50, 52, 53]. Studies have shown that the evoked sense of social presence from anthropomorphic agents' social cues affects trust in online platforms [12, 46, 54, 55]. Social presence has been conceptualized as the degree to which users feel the experience that other intelligent beings coexist and interact with them within the digital environment, even when the virtual artifacts are not real-life [28, 56]. This conceptualization has been adopted in some chatbot studies [1, 57]. Relatedly, social presence has also been conceptualized as the degree of social cues afforded through interacting with the technological source as "warm," "personal," and "sociable" [52, 53, 58]. The present paper follows this definition following most chatbot studies adopting it [11, 59–61].

According to the multiple source effect theory, the feeling of social presence sensed by a user in a digital environment will be amplified when numerous voice sources are embedded within the digital environment [62]. The theory

asserts that as each voice source equals a source of social presence; hence, the presence of multiple voice sources should contribute to a higher sense of perceived social presence. This effect was first shown in a study with a book review website; the implementation of multiple artificial voices led listeners to experience higher social presence than the implementation of a single artificial voice [62]. In a multi-agent study [28], domain-specific IoT devices with multiple unique voices for conveying information about weather, traffic, or event, respectively, led to a higher sense of social presence than the generalist IoT device with a single voice for conveying information about weather, traffic, or event. The finding consequently inspired other scholars [12] to reproduce the multiple source effect with embodied virtual agents. Accordingly, users engaging with multiple virtual agents in a website interface reported a greater social presence than those in the single-agent website interface.

2.5. Users' Gender and Social Cues of Agents. Females and males tend to engage in different processing — while females tend to utilize a peripheral route that focuses on the message source's affective attributes, males tend to use a central route emphasizing the message content [13, 30]. Relatedly, females have superior processing of socio-emotive cues during social interaction in a technologically mediated environment [63]. Hence, they are better at decoding social cues exhibited by anthropomorphic agents and are more likely to form judgments and attitudes based on these cues [64]. A study [30] showed that females tended to react negatively toward multiple female agent voices while responding more approvingly toward a single consistent female agent voice. Conversely, males preferred the multi-agent scenario where unique female voice agents were assigned to different devices. On the other hand, another experiment [13] revealed that while both males and females reacted positively toward the multi-agent platform, the positive effects of specialty cues derived from the multi-agent interface design were more pronounced for females than males. The lack of studies and contrasting findings concerning gender role in a multi-agent context inspires this study to examine the moderating role of user's gender in the effects of multi-chatbot and single-chatbot interfaces.

3. Methodology

3.1. Conceptual Framework. This study's independent variable is the m-commerce chatbot interface type — the multi-chatbot and single-chatbot. Drawing on the literature review, the dependent variables of this study are perceived agent expertise, m-commerce trusting beliefs (ability, benevolence, integrity), intention to purchase through the m-commerce platform, and perceived social presence in the m-commerce platform. Figure 1 illustrates the conceptual framework of this study.

3.2. Stimuli

3.2.1. Chatbot Design. We designed four unique chatbot avatars resembling females in graphical cartoon format to embody the chatbot personas. The agents were gendered as

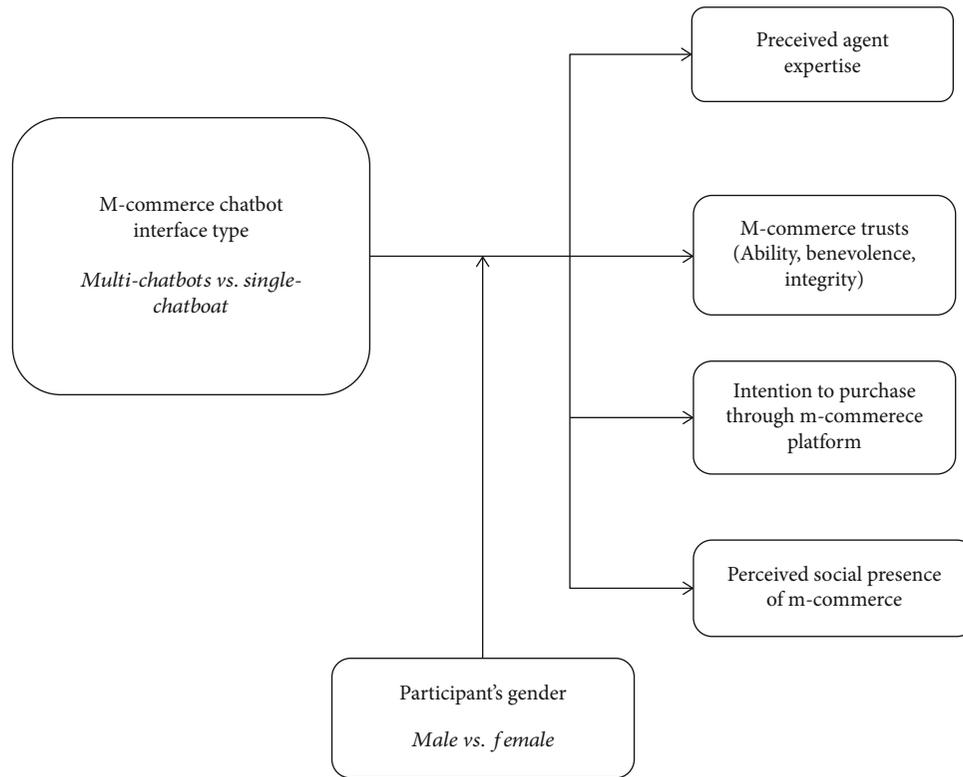


FIGURE 1: Conceptual framework of this study.

females as they convey qualities crucial in human-agent interaction such as caring, sincere, and emphatic [43]. Moreover, this study considered cartoonish than a realistic form more appropriate for chatbot avatars in practical settings. In light of the frustration and disappointment that may arise because the current agent technology has yet to match users' expectations fully, avatars rendered in cartoonish semblance can help minimize unrealistic expectations toward the chatbot systems [59, 61, 65–68]. Figure 2 illustrates the chatbot avatars representing the multi-chatbot and single-chatbot interfaces.

3.2.2. M-Commerce Chatbot Platform. The avatars were embedded into a mock interface resembling an m-commerce platform integrated within a Facebook messenger application. With over 300,000 bots as of 2018 [69], the Facebook messenger is one of the most used and familiar messaging applications among users and has been examined in recent user-chatbot interaction research [2, 70]. The mock m-commerce platform offered three product categories: electronics, health products, and home appliances.

3.2.3. Multi-Chatbot and Single-Chatbot Interface Design. Two videos, one for single-chatbot interface (duration = 375 s) and the other for multi-chatbot interface (duration = 380 s), were developed to simulate user-chatbot text conversations within the mock m-commerce platform. The use of video as stimuli was similar to the experimental approach found in prior studies examining the effects of agent design (virtual agent and chatbot) on user perceptions [71, 72].

The simulated conversation involved a user (which the participants would need to imagine themselves as) seeking recommendations for a laptop, a fitness tracker, and a portable air cooler; consequently linking the user's queries to the three product categories offered in the m-commerce platform — electronics, health, and home appliance. These products were considered high-involvement because the purchase decision mirrored one's well-being facets, including health, comfort, and productivity [12, 13]. Moreover, expensive and sophisticated products tend to increase consumers' involvement in purchase decisions as they would have to endure the consequences of purchase risk like defective or mismatched product price and value [73].

The single-chatbot interface video simulated the textual interaction between a user and one chatbot (Lydia) assigned to provide recommendations across all three product categories. The user requested suggestions for a laptop (electronic product zone), then a fitness tracker (health product zone), and finally a portable air cooler (home appliance product zone). Only one chatbot (Lydia) dispensed recommendations for all the products mentioned earlier. The multi-chatbot interface featured a textual conversation between the user and four chatbots — one designated as a storefront manager (Lydia); whereas the other three were appointed as product-specific advisors offering recommendations within their respective product categories, including Chloe, the designated chatbot for electronic product zone; Eva the designated chatbot for health product zone; and Jenny the designated chatbot for home appliance product zone. Therefore, each product-related query by the user was answered



FIGURE 2: The avatars representing the chatbot personas — Lydia the generalist/store manager (L), Chloe the advisor for electronic devices, Jenny the advisor for health products, and Eva the advisor for home appliances (R).

by the respective chatbot assigned to her unique product zone. Moreover, the three specially designated chatbots were set with dialogue scripts that introduced themselves as product-specific advisors (“Greetings to you, I am your chatbot advisor who is dedicated to the electronic devices product category”). In addition, the chatbots were given scripts that introduced fellow chatbots as product-specific advisors, including “Let me introduce Chloe, our chatbot product advisor especially for electronic devices, to advice on your preference.” Apart from the dialogues, the scripts about the product information were constant between the generalist chatbot interface design and domain chatbot interface design.

The videos were reviewed by three academics whose research fields were related to human-computer interaction. Minor adjustments were made to improve the video speed and font visibility, per the acquired feedback. Figure 3 illustrates the videos simulating the mock user-chatbot conversations for the multi-chatbot and single-chatbot interfaces.

3.3. Dependent Measures. The questionnaires used in this study were adapted from relevant studies. The items and instructions were all worded in English, as the participants of this study were pooled from undergraduates at a large private university that uses English as the medium of instruction.

3.3.1. Perceived Competence of Chatbot. A five-point Likert scale asked participants to rate the following statements, including “The chatbot(s) is/are expert(s) on the products,” “The chatbot(s) specialize(s) in the products,” and “The chatbot(s) is/are knowledgeable about the products.” These items were adapted from studies that have assessed virtual agents’ perceived expertise in multi-agent interfaces [12, 13, 27].

3.3.2. Perceived Trust Ability toward M-Commerce Platform. A five-point Likert scale asked participants to rate the following statements, including “This m-commerce platform is competent and effective in its interactions with me,” “This m-commerce platform performs all of its roles very well,” “Overall, this m-commerce platform is capable and proficient,” and “In general, this m-commerce platform is informative.” This scale was adapted from related studies [12,

13, 26], which measured the perceived ability of embodied virtual agents and chatbots in multi-agent technological sources.

3.3.3. Perceived Trust Benevolence toward M-Commerce Platform. Adapted from the scale used in prior works concerning multi-agent interfaces with embodied virtual agents and chatbots [12, 13, 26], the five-point Likert scale asked participants to rate the following statements, including “I believe that this m-commerce platform would act in my best interest,” “If I required help, this m-commerce platform would do its best to provide assistance,” and “This m-commerce platform is interested in my well-being and not just its own.”

3.3.4. Perceived Trust Integrity toward M-Commerce Platform. Adapted from the scale used in prior studies on embodied virtual agents and chatbots in multi-agent interfaces [12, 13, 26], the five-point Likert scale asked participants to rate the following statements, such as “This m-commerce platform is truthful in its dealing with me,” “I would characterize this m-commerce platform website as being honest,” “This m-commerce platform keep its commitments,” and “This m-commerce platform is sincere and genuine.”

3.3.5. Perceived Social Presence in M-Commerce Platform. Adapted from the scale used in research on multi-agent interfaces [12, 58], the measure asked participants to rate on a five-point Likert scale on the following questions, such as “I can feel a sense of human contact with the m-commerce platform,” “I can feel a sense of human sociability with the m-commerce platform,” “I can feel a sense of human warmth with the m-commerce platform,” and “I can feel a sense of human sensitivity with the m-commerce platform.”

3.3.6. Purchase Intention through M-Commerce Platform. Similar to the assessment utilized in the studies on embodied virtual agents and chatbots in multi-agent interfaces [12, 13, 26], the five-point Likert scale contained the following items to be rated by participants, including “I will consider buying a product through the m-commerce platform,” “I would like to try a product from the m-commerce platform,” and “I will surely buy a product from the m-commerce platform.”

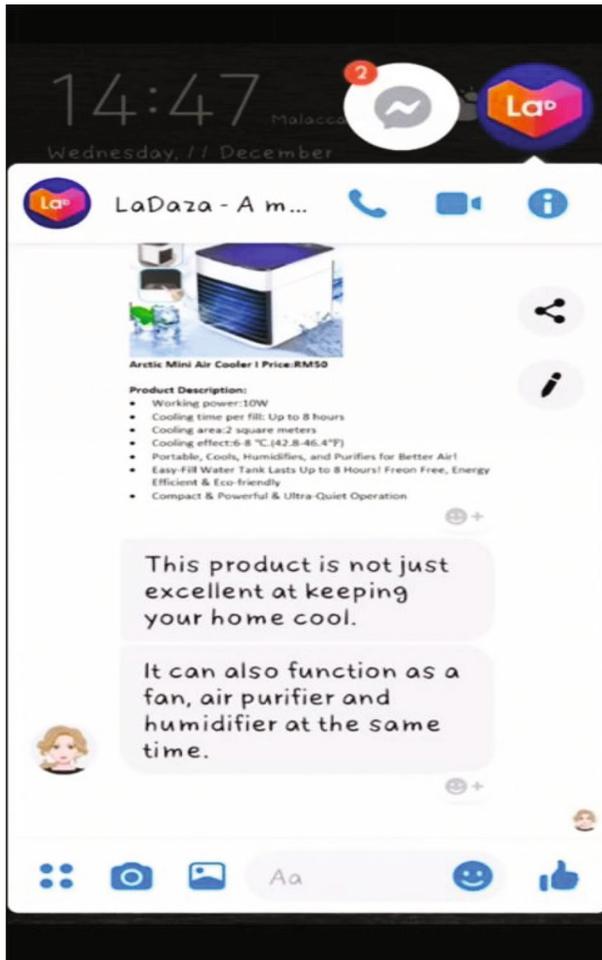


FIGURE 3: Screen shot of simulated user-chatbot interaction.

3.4. Pilot Study Assessing Reliability of Dependent Measures.

A pilot study was conducted to check the reliability of the dependent measures. Thirty-three recruited participants (none of whom participated in the main experiment) watched the stimuli videos and after that filled up the survey set measuring perceived competence of chatbot, perceived ability of m-commerce platform, perceived benevolence of m-commerce platform, perceived integrity of m-commerce platform, perceived social presence, and purchase intention through the m-commerce platform. Reliability analyses showed that all variables had α values of more than 0.7, indicating that the dependent measures were reliable with good and acceptable internal consistency.

3.5. Research Design and Online Experiment. This study employed a between-subjects online experimental design wherein participants were given exposure to either the multi-chatbot or single-chatbot interface design video. Due to the current restrictions to physical meetings, the experiment was conducted using Alchemer, an online survey platform. We inserted a script in the survey platform informing participants that they were about to view a video simulating textual user-chatbot conversation within a multi-product category platform. The script also reminded participants to

imagine themselves as the user in the simulated interaction during the video presentation and that they would need to fill out a survey after the video display.

The online survey tool afforded systematic A/B randomization, and hence, participants were randomly presented with either the multi-chatbot or single-chatbot interface design video. After the video presentation, the survey platform prompted the participants to answer the items measuring perceived competence of chatbot, perceived ability of m-commerce platform, perceived benevolence of m-commerce platform, perceived integrity of m-commerce platform, perceived social presence, and purchase intention through the m-commerce platform.

The authors of this study initiated a word-of-mouth invitation to undergraduates of business and information science major at a large Asian university that uses English as the medium of instruction. We set the online survey tool to disallow mobile device access (such as smartphones and tablets) to ensure that the videos must be viewed with the appropriate frame size. Further, the online tool was set to ensure that the response activity for the survey could only be performed once for each computer to circumvent duplication or repeated attempts. Participants assessed the online survey platform through a given link over four days, yielding a total of 192 completed responses.

3.6. Data Analyses

3.6.1. Filter. Based on two exclusion criteria, we dismissed participants' data who might have been inattentive or experienced technical issues during the online survey. First, we discarded participant's data with incorrect answers to any of the two post-video questions asked in the survey platform: 1) State the number of the product(s) offered in the m-commerce platform, and 2) State the number of the chatbot(s) featured in the m-commerce platform. Further, based on the timestamp records, we dropped participants' data with a video page period less than the respective video duration (indicating that the videos were not viewed fully or at all). Thirty-eight participants' data were removed from the 192 completed responses. Thus, the subsequent data analyses were conducted with 154 valid responses.

3.6.2. Participant's Profile. The data indicated that the participants were all aged between 20 and 25 years old, and the gender ratio was 46% male and 54% female. Concerning e-commerce experience, 32.5% of the participants had less than a year of e-commerce experience, 43.5% had e-commerce experience between 1 year and five years, and 24% had more than five years of e-commerce experience.

3.6.3. Descriptive Data. Table 1 illustrates the means and standard deviations of the dependent variables between the multi-chatbot interface and the single-chatbot interface.

3.6.4. Main Analyses. We conducted a series of 2 (participant's gender: male vs. female) X 2 (m-commerce chatbot interface: multi- chatbot vs. single chatbot) ANOVAs for perceived chatbot competence, social presence, and purchase intention, respectively. Relatedly, a 2 (participant's gender:

TABLE 1: Means and standard deviations of participants between the multi-chatbot interface and the single-chatbot interface.

	Multi-chatbot interface			Single-chatbot interface		
	Male (n = 38)	Female (n = 45)	Total (n = 83)	Male (n = 33)	Female (n = 38)	Total (n = 71)
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Perceived chatbot competence	3.93 (.65)	4.25 (.55)	4.11 (.62)	4.29 (.55)	4.09 (.56)	4.19 (.56)
Platform trust ability	3.81 (.69)	4.11 (.57)	3.97 (.64)	4.41 (.53)	4.32 (.54)	4.36 (.54)
Platform trust benevolence	3.65 (.71)	3.98 (.79)	3.82 (.77)	4.03 (.69)	4.11 (.72)	4.07 (.70)
Platform trust integrity	3.77 (.69)	4.05 (.67)	3.92 (.69)	4.23 (.57)	4.09 (.66)	4.16 (.62)
Intention to purchase through platform	3.77 (.68)	4.11 (.72)	3.95 (.71)	4.15 (.78)	3.88 (.89)	4.01 (.85)
Perceived social presence	3.61 (.71)	3.91 (.84)	3.77 (.79)	4.06 (.62)	3.99 (.55)	4.02 (.58)

male vs. female) X 2 (m-commerce chatbot interface: multi-chatbot vs. single chatbot) MANOVA was also conducted for m-commerce platform trust ability, m-commerce platform trust benevolence, and m-commerce platform trust integrity; given that the variables were conceptually related. For the two-way ANOVAs and the planned comparisons through one-way ANOVAs, Levene's test for homogeneity of variances for each analysis showed that the significance values were greater than .05; hence, the homogeneity of variance assumptions was not violated. Likewise, the two-way MANOVA indicated the values produced by Box's Test of Equality of Covariance Matrices as $p > .001$, as well as Levene's test for homogeneity of variance for each dependent variable as $p > .05$. These values, therefore, affirmed that the assumption of homogeneity of variance-covariance matrices and the premise of the equality of variance for each variable were not violated.

3.6.5. Perceived Chatbot Competence. Concerning perceived chatbot competence, the result found significant interaction effect between the chatbot interface types and the user's gender on the dependent variable [$F(1,150)=7.64$, $p = .00$, $\eta^2 = .05$]. Additional one-way ANOVAs split by the user's gender demonstrated that for male participants particularly, the single-chatbot m-commerce interface led to significantly higher perceived chatbot competence than the multi-chatbot m-commerce interface [$F(1,69)=6.39$, $p = .01$, $\eta^2 = .09$]. The main effect of the chatbot interface types on perceived chatbot competence was non-significant [$F(1,150)=1.25$, $p = .52$, $\eta^2 = .008$].

3.6.6. Trusting Beliefs toward the M-Commerce Platform. The two-way MANOVA for trusting beliefs toward the m-commerce platform (ability, benevolence, and integrity) indicated a significant difference between the chatbot interface types on the combined dependent variables [$F(3,148)=6.31$, $\Lambda = .887$, $p = .00$, $\eta^2 = .113$]. This qualifies further examination to uncover the between-subjects effects for each dependent variable. As per the Bonferroni adjustment of alpha level with three dependent variables, we considered the between-subjects effects significant at $p = .017$ [74]. The results revealed a significant main effect of the m-commerce chatbot interface types on platform trust ability [$F(1,150)=17.87$, $p = .00$, $\eta^2 = .106$] — specifically, participants attributed significantly higher trust ability toward the single-chatbot m-commerce interface than the multi-

chatbot m-commerce interface. While not reaching statistical significance, participants in the single-chatbot m-commerce interface condition reported marginally higher trust benevolence toward the m-commerce platform [$F(1,150)=4.615$, $p = .033$, $\eta^2 = .03$] as well as higher trust integrity toward the m-commerce platform [$F(1,150)=5.702$, $p = .018$, $\eta^2 = .037$] than those in the multi-chatbot m-commerce interface condition.

3.6.7. Purchase Intention through the M-Commerce Platform. With respect to purchase intention through the m-commerce platform, the result found significant interaction effect between the chatbot interface types and the user's gender on the dependent variable [$F(1,150)=5.84$, $p = .01$, $\eta^2 = .04$]. Additional one-way ANOVAs split by the user's gender demonstrated that for male participants particularly, the single-chatbot m-commerce interface led to significantly higher purchase intention through the m-commerce platform than the multi-chatbot m-commerce interface design [$F(1,69) = .481$, $p = .03$, $\eta^2 = .06$].

3.6.8. Perceived Social Presence. Concerning perceived social presence, the result indicated a significant main effect of m-commerce chatbot interface types on perceived social presence [$F(1,150) = 5.70$, $p = .02$, $\eta^2 = .04$] — specifically, participants who interacted with the single-chatbot m-commerce interface reported significantly higher perceived social presence than those who interacted with the multi-chatbot m-commerce interface. The interaction effect between the chatbot interface types and the user's gender was non-significant [$F(1,150) = 2.78$, $p = .098$, $\eta^2 = .018$].

3.7. Mediation Analyses. The main analyses indicated that the male participants ascribed significantly higher perceived chatbot competence when they engaged with the single-chatbot than the multi-chatbot interface. Based on the literature demonstrating the association between source credibility of chatbot/agent and purchase intention [11, 12, 26, 43], we were interested in determining perceived chatbot competence as a mediator in the relationship between chatbot interface types (multi-chatbot vs. single-chatbot) and intention to purchase through the platform among the male participants. The PROCESS-macro mediation analysis using the indirect bootstrapping method with 5000 bootstrap resamples (Preacher & Hayes, 2004) affirmed the mediating role of perceived chatbot competence in the effects of chatbot

interface types on intention to purchase through the platform among the male participants, $B = -.16$, 95% CI $[-.36, -.03]$. Furthermore, when controlling for perceived chatbot competence, the correlation between chatbot interface types and intention to purchase through the platform became non-significant, $B = -.21$, $t(69) = -1.28$, $p = .20$; thus indicating that perceived chatbot competence was a full mediator in the relationship for the male participants. Taken together, for the male participants, the single-chatbot interface (versus the multi-chatbot interface) enhanced perceived chatbot competence, which in turn led to higher intention to purchase through the platform.

The previous comparison analyses revealed that the single-chatbot interface led to significantly higher perceived trust ability of the platform and a significantly higher sense of social presence for all participants in general. Given these findings, we next explored whether perceived social presence mediated the effects of chatbot interface types on perceived trust ability of the platform, based on the literature showing that social presence from anthropomorphic agents can influence trusting beliefs toward the online platforms [12, 46, 53, 54]. The PROCESS-macro mediation analysis using the indirect bootstrapping method with 5000 bootstrap resamples endorsed the mediating role of perceived social presence in the effects of chatbot interface types on the trust ability of the platform, $B = -.09$, 95% CI $[-.21, -.02]$. When controlling for perceived social presence, the correlation between the chatbot interface types and trust ability remained significant, $B = -.29$, $t(152) = -3.34$, $p = .00$; thus indicating that perceived social presence was a partial mediator in the relationship. In other words, the single-chatbot interface (versus the multi-chatbot interface) induced a stronger sense of social presence in the platform, which partially contributed to higher trust toward the platform's ability among the participants. Table 2 summarizes this study's findings.

4. Discussion

This study revealed that participants attributed a stronger sense of social presence and trusting beliefs toward the m-commerce platform when engaging with the single-chatbot interface than the multi-chatbot interface. Particularly for males, the single-chatbot interface (vs. the multi-chatbot interface) also led to enhanced perception of chatbot competence and higher intention to purchase through the m-commerce platform. Framing chatbots as domain-specific product advisors in the multi-chatbot interface failed to trigger appropriate specialization schema in participants' minds and hence did not evoke participants to categorize the chatbots as specialists. Thus, this study, together with prior research [14, 30, 31], challenges the robustness of the single functionality as a specialty cue in evoking users to categorize and perceive domain-specific agents as specialists.

A finer-grained analysis of the literature revealed some inferences surrounding this observation. In particular, the single functionality as specialty cue effect can be reinforced with authoritative labels/social descriptors accentuating specialization. Per the categorization theory [75], users across two experiments attributed greater perceived agent compe-

tence when social descriptors signaling specialization ("Hi, I am Anna — your camera product specialist") were affixed to the product-specific agents [12, 13]. However, this study did not devise the product-specific chatbots in the multi-chatbot interface to present themselves as a "product specialist" or "product expert." Instead, the chatbots in the multi-chatbot interface were scripted to introduce themselves as "dedicated product advisor for [product category]," which might not have conveyed a strong enough cue for evoking the participants to categorize them as products specialists. Hence, the mere assignment of chatbots to advise on specific product domains the multi-chatbot interface, albeit devoid of authoritative labels/social descriptors signaling specialization such as "product specialist" or "product expert," was ineffective in producing a social cue triggering specialization schema in user's mind. This could account for the lack of positive effects of single functionality as specialty cues in multi-agent studies incorporating domain-specific agents without labels/social descriptors accentuating specialization [30, 31]. Specifically, researchers in a similar study [30] implied that the labels affixed to the voice agents, i.e., "SIRI" and "SORI," were novel to users and held no cues for prompting users to categorize the entities as specialists or generalists.

Why did the single-chatbot m-commerce interface evoke higher perceived chatbot competence and greater intention to purchase through the m-commerce platform, particularly among the male participants? The mere assignment of chatbots to advise on specific product domains without robust labels/social descriptors highlighting specialization could have prompted the male participants to appraise these chatbots as novice product advisers possessing limited product domain knowledge. On the other hand, the generalist chatbot in the single-chatbot interface was socially perceived by the male participants as more competent, based on her broader expertise in dispensing advice across diverse product categories. Per the mediation analyses, the higher perceived chatbot competence induced by the single-chatbot interface (vs. multi-chatbot interface) led to higher purchase intention through the platform among the males. This observation supports the notion that the source credibility of agents (agent's competence or expertise) can drive the user's intention to use the agent-infused platforms [11, 13, 27, 43].

Why did the single-chatbot interface evoke higher trusting beliefs (ability, benevolence, and integrity) toward the m-commerce platform than the multi-chatbot interface? The mere stratagem of designating chatbots to advise on specific product categories (without labels as specialty cues) was ineffective in invoking specialization heuristics in participants' minds; hence, the multi-chatbot interface did not convey positive social cues for accentuating source credibility or the reputation of the m-commerce platform. Without source credibility cues to function as "proof sources," participants could not easily verify the expertise and the trustworthiness of the multi-chatbot m-commerce platform [11, 12, 26, 41, 43], thereby causing them to attribute less trusting beliefs toward the multi-chatbot m-commerce platform. Recent studies have found that users felt more confusion [15] and

TABLE 2: Summary of research findings.

	Findings
Perceived chatbot competence	Male participants who engaged with the single-chatbot interface attributed the chatbot with higher perceived competence than the male participants who engaged with the multi-chatbot interface.
Trust toward m-commerce platform	Participants who engaged with the single-chatbot interface attributed higher trust toward the m-commerce platform (ability, benevolence, and integrity) than those who engaged with the multi-chatbot interface.
Purchase intention through the m-commerce platform	Male participants who engaged with the single-chatbot interface reported higher purchase intention through the m-commerce platform than the male participants who engaged with the multi-chatbot interface.
Perceived social presence	Participants who engaged with the single-chatbot interface reported higher social presence in the m-commerce platform than those who engaged with the multi-chatbot interface.
Mediation effects	For all participants, perceived social presence was a partial mediator in the effects of chatbot interface (multi-chatbot vs. single-chatbot) on trust ability toward m-commerce platform. For males, perceived chatbot competence was a partial mediator in the effects of chatbot interface (multi-chatbot vs. single-chatbot) on intention to purchase through the m-commerce platform.

favor interacting with a consistent agent than many agents as multiple sources can impose extraneous cognitive load [30]. Further, the participants could have sensed the multi-chatbot interface as unfamiliar and not the market’s standard interface. The unfamiliarity of the multi-chatbot interface was compounded by the participants experiencing confusion and added cognitive load when attending to the different chatbots. In this regard, the unfamiliarity cues above could have harmed trusting beliefs toward the multi-chatbot m-commerce platform.

Grounded on the multiple source effect theory [62], studies have shown that multiple voice sources can induce a stronger sense of social presence in users than single-voice interfaces [12, 13, 28]. This study demonstrated a reversal effect on the perceived social presence (single-chatbot leads to higher social presence than multi-chatbot). Hence, this study did not support the multiple source effects with text-cued chatbots. Plausibly, the single-chatbot interface (versus the multi-chatbot interface) led to a higher sense of social presence because the single agent following users across diverse domains naturally affords a longer user-agent interaction duration than a multi-agent interface where each of the many agents has a more brisk interaction with users [31]. The single-chatbot interface might have allowed the participants to experience a longer duration of user-chatbot interaction because the same chatbot was present throughout the recommendation process across the many product categories, thereby eliciting the perceptions that the m-commerce platform was imbued with human-like social qualities. Conversely, the participants engaged with the multi-chatbot interface had inadequate time to effectively sense a social presence in the platform afforded by each chatbot — as the user-chatbot conversation cascaded from one chatbot to another rapidly and transiently. The mediation analyses revealed that the enhanced social presence felt by the participants in the single-chatbot interface condition partially explained the effects of the single-chatbot interface (vs. the multi-chatbot interface) on increased trust ability toward the m-commerce platform. To some degree, this aligns with the extant literature show-

ing that the sense of social presence derived from an anthropomorphic agent’s presence and cues is positively associated with the user’s trust toward the agent-infused platform [12, 13, 28, 41, 45, 46, 54, 55].

There were some gender effects in this study. Only males were affected by the different chatbot interfaces concerning perceived chatbot competence and intention to purchase through the platform. This finding could be attributed to the theoretical perspective that male online shoppers emphasize utilitarian and convenience qualities that guarantee to receive practical values; whereas, females focus on trustworthiness and the ability to interact within digital commerce platforms socially [76–78]. Moreover, utilitarian values are often associated with the qualities of the information and interaction afforded by recommender agents as message sources [66, 79, 80]. Possibly, the effects of the chatbot interface type on perceived chatbot competence and intention to purchase through the platform were more pronounced for males than females because they focused more attentively on the chatbot’s source credibility cue to appraise the practical and convenience values of the m-commerce platform, which then drive their decision to patronize the system.

5. Design and Managerial Implications

This study implies that assigning chatbots to specific domains may not be adequate for evoking users to categorize the chatbots as products specialists. Thus, we recommend converging other cues such as labels (“wine specialist”) [29, 81], dialogues, such as “I am Adrienne — a home decoration specialist advisor!” [12, 13, 26], and social attributes including gender or age [25, 43] with the chatbot’s unique product assignment to activate specialization heuristics in users’ minds. Without specialty cues, designers should favor a single-chatbot interface over a multi-chatbot interface because the latter can induce unfamiliarity, confusion, and unnecessary cognitive load to users [14, 15, 30]. A single-chatbot interface that follows users throughout the product recommendation process will evoke firmer trusting beliefs

toward the m-commerce platform and a more significant social presence than the multi-chatbot interface featuring different chatbots as domain-specific agents within the user-chatbot experience.

M-commerce managers and e-tailers should be aware that users react socially to chatbots. In light of this, besides the functional and technical design aspects, chatbot social design should be considered carefully, as social cues can influence users' perceived source credibility, trust, and purchase intention. Moreover, m-commerce managers and e-tailers should note the user's gender differences in m-commerce expectations and experiences; and that males and females may approach user-chatbot engagements differently. Specifically, males may focus on practical and convenience values while females may emphasize social and emotional values in m-commerce. Hence, m-commerce managers and e-tailers can incorporate different chatbot interfaces to conform with the expectations of male and female shoppers.

6. Limitations and Recommendations for Future Research

Some limitations of this study are acknowledged. First, the sample participants of the experiment consisted of undergraduates in an Asian university, thereby limiting the generalizability of findings for a broader group of users. Hence, replication and reproducibility of this research can be conducted to include respondents with different demographic profiles. Second, per the approach used in prior human-agent interaction studies [71, 72], this experiment used videos simulating user-chatbot engagement as stimuli for assessing user's perception and use intention concerning the multi-chatbot interface design and single-chatbot interface design. This endeavor might have produced results that failed to capture nuanced perceptions indicative of users' actual engagement with chatbots in natural settings. To increase the generalizability of experimental findings to authentic user-chatbot engagement, future studies can utilize existing chatbot platforms [14] or Wizard-of-Oz experiments [15] for establishing a more genuine user-chatbot interaction. The present experiment involved transient simulations of user-agent interaction; and, thus, might have captured the user's perception and intention based on initial impressions. Future research can utilize stimuli with extended duration to ascertain the perceptual and behavioral effects of framing chatbots as domain-specific specialist agents that go beyond initial interaction with the entities.

Concerning multi-chatbot or multi-agent interface, future studies can be conducted to tease apart the subtle and cumulative effects of labels/social descriptors and single functionality as specialty cue effects. Based on the proposition that some labels may not produce a strong enough cue for triggering specialization schema, future studies can also examine the effectiveness or "strength" of different parameters of specialty cues, including different terms and phrases in labels. Moreover, the effects of other potential specialty cues apart from derivatives of labels/social descriptors and single functionality should be explored with chat-

bots or agents within a multi-agent interface, including title ("Dr."), gender, dressing style, and communication style [25]. Future work should further address the effects of induced confusion and extraneous cognitive load in multi-agent systems. Given that gender of users has been found to affect user-chatbot/agent experience differently, future works on multi-chatbot or multi-agent interfaces should further examine the role of gender and e-commerce/m-commerce experience as potential moderating factors. Last but not least, the effects of multi-agent interface should be explored with other social, psychological, and communication theories across diverse application domains like health, education, counseling, and finance.

Data Availability

Data is available on request.

Conflicts of Interest

The authors report no conflict of interest.

Acknowledgments

We thank Wei Ming Pang for assisting with the online experiment. This research was supported by the Malaysia Ministry of Higher Education (FRGS/1/2021/SS0/MMU/02/8) titled "A Design Framework of E-Commerce Chatbot with Social Cues to Enhance User-Chatbot Experience in Malaysia."

References

- [1] T. Araujo, "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*, vol. 85, pp. 183–189, 2018.
- [2] E. Van den Broeck, B. Zarouali, and K. Poels, "Chatbot advertising effectiveness: When does the message get through?," *Computers in Human Behavior*, vol. 98, pp. 150–157, 2019.
- [3] D. Zumstein and S. Hundertmark, "Chatbots—an interactive technology for personalized communication, transactions and services," *IADIS International Journal on WWW/Internet*, vol. 15, 2017.
- [4] Z. Kaddari, Y. Mellah, J. Berrich, M. G. Belkasm, and T. Bouchentouf, "Natural Language Processing: Challenges and Future Directions," in *International Conference on Artificial Intelligence & Industrial Applications*, pp. 236–246, Meknes, Morocco, 2021.
- [5] A. K. Tang, "A systematic literature review and analysis on mobile apps in m-commerce: implications for future research," *Electronic Commerce Research and Applications*, vol. 37, article 100885, 2019.
- [6] M. V. Eeuwens, *Mobile conversational commerce: messenger chatbots as the next interface between businesses and consumers*, 2017.

- [7] S. Tuzovic and S. Paluch, "Conversational commerce-a new era for service business development," in *Service Business Development*, pp. 81–100, Springer Gabler, 2018.
- [8] C. H. Leung and W. T. Yan Chan, "Retail chatbots: the challenges and opportunities of conversational commerce," *Journal of Digital & Social Media Marketing*, vol. 8, no. 1, pp. 68–84, 2020.
- [9] H. Clare, *80% of customers don't trust chatbots for aftersales advice*, 2018, <http://www.itpro.co.uk/machine-learning/30606/80-of-customers-dont-trust-chatbots-for-aftersales-advice>.
- [10] A. P. Chaves and M. A. Gerosa, "How Should My Chatbot Interact? A Survey on Social Characteristics in Human-Chatbot Interaction Design," *International Journal of Human-Computer Interaction*, vol. 37, no. 8, pp. 729–758, 2021.
- [11] C. Yen and M.-C. Chiang, "Trust me, if you can: a study on the factors that influence consumers' purchase intention triggered by chatbots based on brain image evidence and self-reported assessments," *Behaviour & Information Technology*, vol. 40, no. 11, pp. 1177–1194, 2021.
- [12] S.-M. Tan and T. W. Liew, "Designing Embodied Virtual Agents as Product Specialists in a Multi-Product Category E-Commerce: The Roles of Source Credibility and Social Presence," *International Journal of Human-Computer Interaction*, vol. 36, no. 12, pp. 1136–1149, 2020.
- [13] T. W. Liew and S.-M. Tan, "Exploring the effects of specialist versus generalist embodied virtual agents in a multi-product category online store," *Telematics and Informatics*, vol. 35, no. 1, pp. 122–135, 2018.
- [14] C. S. Pinhanez, H. Candello, M. C. Pichiliani et al., "Different but equal: comparing user collaboration with digital personal assistants vs. teams of expert agents," 2018, <https://arxiv.org/abs/1808.08157>.
- [15] A. P. Chaves and M. A. Gerosa, "Single or Multiple Conversational Agents? An Interactional Coherence Comparison," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, Montreal, QC, Canada, 2018.
- [16] R. B. Kantharaju and C. Pelachaud, "Towards Developing a Model to Handle Multiparty Conversations for Healthcare Agents," in *Proceedings of the AAMAS Workshop on Intelligent Conversation Agents in Home and Geriatric Care Applications*, pp. 30–34, Stockholm, Sweden, 2018, <http://ceur-ws.org/Vol-2338/paper4.pdf>.
- [17] J. F. Zolitschka, "A Novel Multi-Agent-Based Chatbot Approach to Orchestrate Conversational Assistants," in *International Conference on Business Information Systems*, pp. 103–117, Colorado Springs, USA, 2020.
- [18] D. Wang and H. Fang, "An Adaptive Response Matching Network for Ranking Multi-Turn Chatbot Responses," in *International Conference on Applications of Natural Language to Information Systems*, pp. 239–251, Saarbrücken, Germany, 2020.
- [19] H. Candello and C. Pinhanez, "Recovering from dialogue failures using multiple agents in wealth management advice," in *Studies in Conversational UX Design*, pp. 139–157, Springer, 2018.
- [20] H. Candello, M. Vasconcelos, and C. Pinhanez, "Evaluating the conversation flow and content quality of a multi-bot conversational system," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, Denver Colorado, USA, 2017.
- [21] J. Kim, O. Shinhyeok, O.-W. Kwon, and H. Kim, "Multi-Turn Chatbot Based on Query-Context Attentions and Dual Wasserstein Generative Adversarial Networks," *Applied Sciences*, vol. 9, no. 18, pp. 3908–3908, 2019.
- [22] C. Nass and J. Steuer, "Voices, Boxes, and Sources of Messages," *Human Communication Research*, vol. 19, no. 4, pp. 504–527, 1993.
- [23] C. Nass and Y. Moon, "Machines and Mindlessness: Social Responses to Computers," *Journal of Social Issues*, vol. 56, no. 1, pp. 81–103, 2000.
- [24] J. Feine, U. Gnewuch, S. Morana, and A. Maedche, "A Taxonomy of Social Cues for Conversational Agents," *International Journal of Human-Computer Studies*, vol. 132, pp. 138–161, 2019.
- [25] T. W. Liew and S. M. Tan, "Social cues and implications for designing expert and competent artificial agents: a systematic review," *Telematics and Informatics*, vol. 65, p. 101721, 2021.
- [26] T. W. Liew, S. M. Tan, J. Tee, and G. G. G. Goh, "The effects of designing conversational commerce chatbots with expertise cues," in *2021 14th International Conference on Human System Interaction (HSI)*, pp. 1–6, Gdańsk, Poland, 2021.
- [27] K. J. Kim, "Can smartphones be specialists? Effects of specialization in mobile advertising," *Telematics and Informatics*, vol. 31, no. 4, pp. 640–647, 2014.
- [28] K. J. Kim, "Interacting socially with the internet of things (IoT): effects of source attribution and specialization in human-IoT interaction," *Journal of Computer-Mediated Communication*, vol. 21, no. 6, pp. 420–435, 2016.
- [29] Y. J. Koh and S. S. Sundar, "Effects of specialization in computers, web sites, and web agents on e-commerce trust," *International Journal of Human-Computer Studies*, vol. 68, no. 12, pp. 899–912, 2010.
- [30] S. Jung, K. M. Lee, and F. Biocca, "Voice Control System and Multiplatform Use: Specialist Vs. Generalist," in *International Conference on Human Interface and the Management of Information*, pp. 607–616, Los Angeles, CA, USA, 2015.
- [31] Y. Moon, K. J. Kim, and D. H. Shin, "Voices of the Internet of Things: An Exploration of Multiple Voice Effects in Smart Homes," in *International Conference on Distributed, Ambient, and Pervasive Interactions*, pp. 270–278, Toronto, Canada, 2016.
- [32] C. I. Hovland and W. Weiss, "The Influence of Source Credibility on Communication Effectiveness," *Public Opinion Quarterly*, vol. 15, no. 4, pp. 635–635, 1951.
- [33] B. J. Fogg and H. Tseng, "The elements of computer credibility," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pp. 80–87, Pittsburgh, Pennsylvania, 1999.
- [34] S. Tseng and B. J. Fogg, "Credibility and computing technology," *Communications of the ACM*, vol. 42, no. 5, pp. 39–44, 1999.
- [35] R. Ohanian, "Construction and Validation of a Scale to Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness," *Journal of Advertising*, vol. 19, no. 3, pp. 39–52, 1990.
- [36] C. Pornpitakpan, "The Effect of Celebrity Endorsers' Perceived Credibility on Product Purchase Intention," *Journal of International Consumer Marketing*, vol. 16, no. 2, pp. 55–74, 2004.
- [37] K. H. Yoo and U. Gretzel, "Creating more credible and persuasive recommender systems: the influence of source

- characteristics on recommender system evaluations,” in *Recommender systems handbook*, pp. 455–477, Springer, 2011.
- [38] J. J. Dijkstra, W. B. G. Liebrand, and E. Timminga, “Persuasiveness of expert systems,” *Behaviour & Information Technology*, vol. 17, no. 3, pp. 155–163, 1998.
- [39] J. J. Jiang, G. Klein, and R. G. Vedder, “Persuasive expert systems: the influence of confidence and discrepancy,” *Computers in Human Behavior*, vol. 16, no. 2, pp. 99–109, 2000.
- [40] B. J. Fogg, “Persuasive technology,” *Ubiquity*, vol. 2002, no. -December, p. 5, 2002.
- [41] B. Xiao and I. Benbasat, “E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact,” *MIS Quarterly*, vol. 31, no. 1, p. 137, 2007.
- [42] P. M. Doney and J. P. Cannon, “An examination of the nature of trust in buyer-seller relationships,” *Journal of Marketing*, vol. 61, no. 2, pp. 35–51, 1997.
- [43] A. Beldad, S. Hegner, and J. Hoppen, “The effect of virtual sales agent (VSA) gender - product gender congruence on product advice credibility, trust in VSA and online vendor, and purchase intention,” *Computers in Human Behavior*, vol. 60, pp. 62–72, 2016.
- [44] D. Harrison McKnight, V. Choudhury, and C. Kacmar, “The impact of initial consumer trust on intentions to transact with a web site: a trust building model,” *The journal of strategic information systems*, vol. 11, no. 3-4, pp. 297–323, 2002.
- [45] I. Benbasat and W. Wang, “Trust In and Adoption of Online Recommendation Agents,” *Journal of the Association for Information Systems*, vol. 6, no. 3, pp. 72–101, 2005.
- [46] V. Chattaraman, W.-S. Kwon, and J. E. Gilbert, “Virtual agents in retail web sites: Benefits of simulated social interaction for older users,” *Computers in Human Behavior*, vol. 28, no. 6, pp. 2055–2066, 2012.
- [47] R. E. Goldsmith, B. A. Lafferty, and S. J. Newell, “The Impact of Corporate Credibility and Celebrity Credibility on Consumer Reaction to Advertisements and Brands,” *Journal of Advertising*, vol. 29, no. 3, pp. 43–54, 2000.
- [48] B. A. Lafferty, R. E. Goldsmith, and S. J. Newell, “The Dual Credibility Model: The Influence of Corporate and Endorser Credibility on Attitudes and Purchase Intentions,” *Journal of Marketing Theory and Practice*, vol. 10, no. 3, pp. 1–11, 2002.
- [49] S. W. Wang, G. H.-Y. Kao, and W. Ngamsiriudom, “Consumers’ attitude of endorser credibility, brand and intention with respect to celebrity endorsement of the airline sector,” *Journal of Air Transport Management*, vol. 60, pp. 10–17, 2017.
- [50] D. Gefen and D. W. Straub, “Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services,” *Omega*, vol. 32, no. 6, pp. 407–424, 2004.
- [51] D. Gefen, “E-commerce: the role of familiarity and trust,” *Omega*, vol. 28, no. 6, pp. 725–737, 2000.
- [52] B. Lu, W. Fan, and M. Zhou, “Social presence, trust, and social commerce purchase intention: An empirical research,” *Computers in Human Behavior*, vol. 56, pp. 225–237, 2016.
- [53] D. Cyr, K. Hassanein, M. Head, and A. Ivanov, “The role of social presence in establishing loyalty in e-service environments,” *Interacting with Computers*, vol. 19, no. 1, pp. 43–56, 2007.
- [54] T. W. Liew, S.-M. Tan, and H. Ismail, “Exploring the effects of a non-interactive talking avatar on social presence, credibility, trust, and patronage intention in an e-commerce website,” *Human-centric Computing and Information Sciences*, vol. 7, no. 1, p. 42, 2017.
- [55] T. Hess, M. Fuller, University of Massachusetts, Amherst, D. Campbell, and Millsaps College, “Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents,” *Journal of the Association for Information Systems*, vol. 10, no. 12, pp. 889–919, 2009.
- [56] F. Biocca, “The cyborg’s dilemma: progressive embodiment in virtual environments [1],” *Journal of computer-mediated communication*, vol. 3, no. 2, pp. 324–324, 1997.
- [57] S. Lee, N. Lee, and Y. J. Sah, “Perceiving a Mind in a Chatbot: Effect of Mind Perception and Social Cues on Co-presence, Closeness, and Intention to Use,” *International Journal of Human-Computer Interaction*, vol. 36, no. 10, pp. 930–940, 2020.
- [58] Y. Yoo and M. Alavi, “Media and Group Cohesion: Relative Influences on Social Presence, Task Participation, and Group Consensus,” *MIS Quarterly*, vol. 25, no. 3, p. 371, 2001.
- [59] E. Go and S. S. Sundar, “Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions,” *Computers in Human Behavior*, vol. 97, pp. 304–316, 2019.
- [60] M. Adam, M. Wessel, and A. Benlian, “AI-based chatbots in customer service and their effects on user compliance,” *Electronic Markets*, vol. 31, no. 2, pp. 427–445, 2021.
- [61] R. De Cicco, S. C. e. Silva, and F. R. Alparone, “Millennials’ attitude toward chatbots: an experimental study in a social relationship perspective,” *International Journal of Retail & Distribution Management*, vol. 48, no. 11, pp. 1213–1233, 2020.
- [62] K. M. Lee and C. Nass, “The multiple source effect and synthesized speech: doubly-disembodied language as a conceptual framework,” *Human Communication Research*, vol. 30, no. 2, pp. 182–207, 2004.
- [63] A. R. Dennis, S. T. Kinney, and Y.-T. C. Hung, “Gender Differences in the Effects of Media Richness,” *Small Group Research*, vol. 30, no. 4, pp. 405–437, 1999.
- [64] L. Qiu and I. Benbasat, “A study of demographic embodiments of product recommendation agents in electronic commerce,” *International Journal of Human-Computer Studies*, vol. 68, no. 10, pp. 669–688, 2010.
- [65] J. T. Luo, P. McGoldrick, S. Beatty, and K. A. Keeling, “On-screen characters: their design and influence on consumer trust,” *Journal of Services Marketing*, vol. 20, no. 2, pp. 112–124, 2006.
- [66] M. S. B. Mimoun, I. Poncin, and M. Garnier, “Case study—Embodied virtual agents: An analysis on reasons for failure,” *Journal of Retailing and Consumer Services*, vol. 19, no. 6, pp. 605–612, 2012.
- [67] M. S. B. Mimoun, I. Poncin, and M. Garnier, “Animated conversational agents and e-consumer productivity: The roles of agents and individual characteristics,” *Information & Management*, vol. 54, no. 5, pp. 545–559, 2017.
- [68] S. M. Tan, T. W. Liew, C. L. Gan, and W. M. Wong, “Visual style of embodied virtual sales agents,” *International Journal of Technology and Human Interaction (IJTHI)*, vol. 17, no. 1, pp. 1–13, 2021.
- [69] K. Johnson, *Facebook Messenger Passes 300,000 Bots*, 2018, <https://venturebeat.com/2018/05/01/facebook-messenger-passes-300000-bots/>.

- [70] B. Zarouali, E. Van den Broeck, M. Walrave, and K. Poels, "Predicting Consumer Responses to a Chatbot on Facebook," *Cyberpsychology, Behavior, and Social Networking*, vol. 21, no. 8, pp. 491–497, 2018.
- [71] L. M. Assink, *Exploring Users' Perception of Chatbots in a Mobile Commerce Environment: Creating a Better User Experience by Implementing Anthropomorphic Visual and Linguistic Chatbot Features*, 2019, http://essay.utwente.nl/78306/1/Bachelorthesis_Communication_Science_Lena_Assink.pdf.
- [72] I. Mull, J. Wyss, E. Moon, and S.-E. Lee, "An exploratory study of using 3D avatars as online salespeople," *Journal of Fashion Marketing and Management: An International Journal*, vol. 19, no. 2, pp. 154–168, 2015.
- [73] V. A. Zeithaml, "Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence," *Journal of Marketing*, vol. 52, no. 3, pp. 2–22, 1988.
- [74] B. G. Tabachnick, L. S. Fidell, and J. B. Ullman, *Using Multivariate Statistics*, vol. 5, Pearson, Boston, MA, 2007.
- [75] B. E. Ashforth and R. H. Humphrey, "The Ubiquity and Potency of Labeling in Organizations," *Organization Science*, vol. 8, no. 1, pp. 43–58, 1997.
- [76] B. Hasan, "Exploring gender differences in online shopping attitude," *Computers in Human Behavior*, vol. 26, no. 4, pp. 597–601, 2010.
- [77] L. Zhou, L. Dai, and D. Zhang, "Online shopping acceptance model—a critical survey of consumer factors in online shopping," *Journal of Electronic Commerce Research*, vol. 8, no. 1, p. 41, 2007.
- [78] O. Sohaib, K. Kang, and M. Nurunnabi, "Gender-Based iTrust in E-Commerce: The Moderating Role of Cognitive Innovativeness," *Sustainability*, vol. 11, no. 1, p. 175, 2019.
- [79] R. Etemad-Sajadi and L. Ghachem, "The impact of hedonic and utilitarian value of online avatars on e-service quality," *Computers in Human Behavior*, vol. 52, pp. 81–86, 2015.
- [80] L. T. Huang, "Exploring utilitarian and hedonic antecedents for adopting information from a recommendation agent and unplanned purchase behaviour," *New Review of Hypermedia and Multimedia*, vol. 22, no. 1-2, pp. 139–165, 2016.
- [81] Y. J. Koh and S. S. Sundar, "Heuristic Versus Systematic Processing of Specialist Versus Generalist Sources in Online Media," *Human Communication Research*, vol. 36, no. 2, pp. 103–124, 2010.