 DOI: 10.15706/jksms.2020.21.5.003

Examining the Effect of Chatbot Gender and Gender Congruence between a Chatbot and a Customer in a Banking Context*

Sanjeed, V K(First Author)
Zenyum (Head of Content & SEO)

Kim, Min Gyung(Corresponding Author)
College of Business Management, Hongik University (Assistant Professor)

Wang, Chen-Ya(Co Author)
Institute of Service Science, National Tsing Hua University, Taiwan (Associate Professor)

챗봇의 성별과 챗봇과 고객 간의 성별 일치 효과에 관한 연구: 은행서비스를 배경으로

Sanjeed, V K(제1저자)
Zenyum (Head of Content & SEO)

김 민 경(교신저자)
홍익대학교 상경대학 상경학과 (조교수)

Wang, Chen-Ya(공동저자)
대만 국립칭화대학교 서비스사이언스 대학 (부교수)

* This paper is based on the Master's thesis of V K Sanjeed completed at National Tsinghua University, Taiwan in 2017 under the direction of Dr. Wang, Chen-Ya.

Abstract

Chatbots, automated text-based conversational programs, are a popular technology enabled by artificial intelligence (AI). Chatbots have been adopted in diverse service areas to enhance business performance by facilitating interactions between humans and machines. This study tests the effect of chatbot gender, and the gender congruence between chatbot and customer, on customer responses in a banking context. Data was collected using a scenario-based online survey with a 2 (chatbot gender: male vs. female) x 2 (congruence of chatbot gender and respondent gender: congruent vs. incongruent) between-subject factorial design. Across two chatbot gender and respondent gender contexts, the results show that a male banking chatbot increases customer satisfaction and brand conceptual fluency significantly more than a female banking chatbot. Study results also show the interaction between chatbot gender and the congruence of chatbot gender and respondent gender influences perceived intimacy and information richness. Our findings are relevant to banks, fintech companies, and other businesses that are rapidly adopting chatbots to engage consumers and enhance branding.

Keywords : chatbot, gender, gender congruence, bank, branding
접수일(2020년 10월 14일), 수정일(2020년 11월 27일), 게재확정일(2020년 12월 09일)

I. Introduction

Artificial Intelligent (AI) chatbots are at the frontier of new platforms used for information search, content delivery, and marketing in businesses. A chatbot is a computer program that can mimic “an intelligent conversation with one or more human users via auditory or textual methods” (Gupta et al., 2015, p.1483). Chatbot technology has been utilized in various fields, including marketing, healthcare, e-commerce, retail, government, education, and finance (Luo et al., 2019; Park, 2017). The market size of Chatbots is expected to be more than $1.34 billion in 2024 (Pise, 2018). With wide applications, chatbot systems have become popular due to their accessibility, cost-effectiveness, ease of use, and capacity to process high-volume data (Alharbi, 2020). Chatbots allow businesses to scale customer interaction and routine services without hiring dedicated service personnel. One of the industries that has seen the most active application of chatbots is the finance industry (Byun and Cho, 2020; Kim et al., 2019). AI improves certain metrics required in finance, including automation, speed of decision making, and communication (Alharbi, 2020), and its usefulness has been increasingly highlighted due to the need for and proliferation of untact services (Byun and Cho, 2020). Consequently, numerous companies have made significant investments into research and development of chatbot technology (Alharbi, 2020). However, research has been focused on the technological and systematical dimensions of chatbots, with a lack of academic discussion regarding the attributes and characteristics of chatbot services and their impact on consumer reactions (Byun and Cho, 2020; Kim et al., 2019).

Many industry sectors have utilized chatbot services for diverse purposes. For instance, banks have adopted chatbots for various tasks, such as checking balances, detecting fraud, receiving customer complaints, and providing customized financial products (Byun and Cho, 2020). Bank customers feel familiar and engaged with virtual chatbot assistants, as these assistants mimic human conversation through voice and text messages (Luo et al., 2019; Przegalinska et al., 2019). Most of the chatbots used in the finance industry have female names, and prior research has shown that gender does influence customer perceptions of financial service providers (Söderberg, 2013). Chatbots have been given humanlike attributes such as gender, and customers may have a tendency to evaluate chatbot functions based on emotional connections as opposed to the actual performance of the chatbot (McDonnell and Baxter, 2019). As such the effects of chatbots' humanlike characteristics, including gender, need to be investigated if developers are to increase the efficacy of chatbots. However, there is a lack of understanding regarding how gender affects chatbot service outcomes (Lin et al., 2019). Particularly, customer preferences in terms of chatbot gender that depend on the customer's gender should be understood (Forlizzi et al., 2007). Given that brands give their chatbots gender-specific and/or gender-neutral names, it is important to evaluate the effects of chatbot gender and the interaction between chatbot gender and respondent gender on customer responses such as customer satisfaction.

Nowadays, chatbots are with unique personalities, and they are used for branding strategies (Pavlus, 2016). The underlying assumption is that by using a different communication style, i.e. bold, irreverent, exciting, responsible, a chat-bot will be perceived as more human-like and less like a robot. This in turn is expected to enhance customer satisfaction and other brand associated evaluations. However, since AI chatbots only gained prominence in 2016 (Dale, 2016), re-search centering on communication styles in the chatbot medium is lacking. Hence, the adoption and use of chatbots in banking needs more attention and requires rigorous research in order to determine how to craft more personalized customer service experiences (Alharbi, 2020). To address these gaps, this study examines the effect of the chatbot gender and the gender congruency between a customer and a chatbot on customers' perception and reaction. We also inves-tigate customer responses across two types of chatbot communication style (responsibility vs. activity) for enhancing the generalizability of the study results. We focus on the banking industry as it is one of the major industries rapidly adopting chatbots to compete with fintechs.

Ⅱ. Literature Review

2.1 Chatbots

Chatbots, a compound word from "chatting" and "robots", are the latest iteration of software applications that can converse with humans using natural language processing (NLP) (Dale, 2016; Kim et al., 2019; Suh and Yoon, 2019). However, applications that can hold a dialog via written language were invented decades before the appearance of smartphones and apps. Eliza, developed in the early 1960s used simple pattern matching and a template-based response mechanism to simulate the conversational style of a psychotherapist (Weizenbaum, 1966). Part of the reason why chatbots are so popular is that they can be easily integrated into popular messenger apps such as Facebook Messenger and Line (Song and Choi, 2020). They are also used in Voice Assistants such as Apple's Siri, Amazon's Alexa, and Microsoft's Cortana.

The benefits of chatbots include their relatively low cost, ease of use, and ability to provide personalized customer service (Przegalinska et al., 2019). The benefits of chatbots have been recognized across diverse fields, and chatbots services are expected to continue growing in the future (Song and Choi, 2020). Due to their significant benefits, chatbots have been used in the fields of education (Ranoliya et al., 2017), retail (Chung et al., 2018), healthcare (Morris et al., 2018), and public services (Park, 2017). Chatbots may have found their greatest utility in financial services, including banking and insurance (Byun and Cho, 2020). One of the first and most famous chatbots is Erica of Bank of America. Responding to both text and voice input, Erica can even provide financial planning advice to customers, including how to reduce their interest spending or upgrade their credit line. In South Korea, various untact banking services have been available since 2015, when deregulation of real-name financial verification took effect. Woori Bank's Wibee Talk, introduced in 2017, was the first banking chatbot in South Korea, and other major banks including Kookmin Bank, KEB Hana Bank, and Shinhan Bank now have their own chatbots. These chatbots assist, in various tasks, providing everything from account and card use status information to financial consulting services based on analysis of the customer's consumption pattern (Kim et al., 2019).

Communications with chatbots is based on a pre-programmed dialogue structure, but customers may perceive those communications as interactive and personalized (Kunze, 2016). Chatbots thus allow businesses to scale customer interaction at a fraction of the cost since there is less need to hire dedicated service personnel. This ability to scale service interaction, coupled with the pervasive use of chat apps has made chatbots the focus of new marketing efforts. As they do so, brands have started to differentiate their chatbots with unique personalities (Pavlus, 2016). This is not surprising given that the design of IT artifacts could influence users' perceptions of socialness if they communicated feelings of social presence and sense of personalization (Wang et al., 2007).

2.2 The Effect of Gender and Gender Congruence

2.2.1 Chatbot Gender and Customer Satisfaction

Chatbot gender plays a significant role in interactions between humans and computers. Consequently, the gender of chatbots should be considered a primary feature in chatbot agent design (Forlizzi et al., 2017). Gender is our main dimension of interest for two reasons. Firstly, most chatbots for banks and fintechs have female names such as Cleo, Nina, Erica (Bank of America), and Emma (OCBC Bank) ("Chatbots: The fintech disruptor," 2017). Non-finance businesses such as technology firms including Apple (Siri), Amazon (Alexa) and Microsoft (Tay, Cortana) also use female names for their chatbots and AI assistants. Secondly, studies show that gender can impact consumers' perceptions of virtual agents and the relationship with them (Niculescu et al., 2010). Söderberg (2013) found financial advisor's gender is related to the customers' perception of advisors credibility and customers' willingness to accept the advice from the advisor. Given that banking chatbots act as substitutes for financial advisors, we are interested in investigating the effect of gender on chatbot service encounters. Customers still think in terms of gender stereotypes when interacting with personified agents such as chatbots and prefer it when agent gender corresponds to stereotypical roles (Forlizzi et al., 2007). For instance, chatbot agents might be more likely to stand in for roles traditionally undertaken by females, such as matchmaker, receptionist, or librarian (Niculescu et al., 2010).

Customer satisfaction is a key metric of service quality particularly in banks, where building relationships with customers is critical to long-term success (Ravald and Grönroos, 1996). Customer satisfaction leads to higher repurchase intentions, word-of-mouth, and loyalty amongst customers (Kim and Kim, 2020; Lee, 2018). Major financial institutions have adopted AI to enhance the efficiency of customer services and increase customer satisfaction (Leong 2015). Studies have demonstrated a positive impact on customer satisfaction through the use of AI digital assistants (Okoro, 2014). The services provided by chatbots are unique and convenient as they are available at any time (Nagarhalli et al., 2020). Chatbots facilitate service encounters and transactions for customers, and the overall service experience builds customer satisfaction. The polite, responsive, and friendly attitude of chatbots and their ability to deliver essential information fulfills the needs and expectations of customers (Verhagen et al., 2014). In consumer behavior research, gender is considered to be a powerful variable. In general, compared to their counterparts, women value teamwork and relationships more (Pradhan et al., 2017); these are two key qualities for chatbots. Most existing banking chatbots are female, and as a result, customers may be more familiar with female chatbots while also preferring them due to the stereotypical qualities they personify, qualities traditionally associated with women. Hence, we propose:

Hypothesis 1: A customer encountering a chatbot with a female name (compared to a male name) will report higher Customer Satisfaction (CS) in a chatbot service interaction.

2.2.2 Chatbot Gender and Brand Evaluation

Brands increasingly leverage chatbots as a branding tool (Torba, 2016). The effect of gender on overall brand evaluation of a bank is thus another dimension we are interested in measuring. Overall brand evaluation measures a customers' affective responses toward a brand including items such as liking, desirability, and trust of the brand (Sirianni et al., 2013). When customers evaluate brands, they use past and new information related to brands, and their reactions manifest as a preference for some brands and an aversion to others (Bapat and Thanigan, 2016). Customers encounter diverse brand experiences by searching, shopping, and using products and services (Brakus et al., 2009). Positive brand experiences lead to positive brand evaluation, and when a brand experience involves emotions, it is more likely to lead to a positive brand evaluation (Bapat and Thanigan, 2016). The competence and emotional characteristics attributed to brands may be affected by the gender with which the brand is associated, hence gender can play a crucial role in brand evaluation (Pradhan et al., 2017). Gender-related characteristics can be very prominent in the minds of consumers (Aaker, 1997). It is thus likely that consumers will take into account the gender of a brand in their evaluation of the brand (Lieven et al., 2014). And so female chatbots, with their stronger emotional appeal and stereotypical association with the role, may be more suitable as financial assistants. Therefore, we propose;

Hypothesis 2: A customer encountering a chatbot with a female name (compared to a male name) will report higher overall Brand Evaluation (BE) in a chatbot service interaction.

2.2.3 Chatbot Gender and Brand Conceptual Fluency

Conceptual fluency is defined as “the ease with which customers can process and understand information such as brand meaning” (Sirianni et al., 2013, p.110). To assist customers in becoming conceptually fluent, a target can be presented in a predictive way, or customers can be primed by means of an associated construct (Sirianni et al., 2013). When consumers experience new stimuli for a certain brand and that stimuli corresponds to the previously experienced brand image or their other knowledge of the brand, brand conceptual fluency is enhanced. In other words, brand conceptual fluency is an outcome of strong associations between a stimulus and a brand. Brands that instill in customers a high level of brand conceptual fluency thus have greater brand appeal and tend to receive higher evaluations from customers (Lee and Labroo, 2004). As a banking chatbot's gender constitutes part of the information presented to customers, it can thus influence brand conceptual fluency. It is also more likely that customers who have been previously exposed to banking chatbot services have encountered female chatbots as opposed to male chatbots. Therefore, banking customers may feel more familiar with female service assistants or advisors, including chatbots, and may expect to encounter female service agents more than their male counterparts. Thus, we propose;

Hypothesis 3: A customer encountering a chatbot with a female name (compared to a male name) will report higher Brand Conceptual Fluency (BCF) in a chatbot service interaction.

2.2.4 Chatbot Gender and Information Richness

One of the major tasks of a chatbot is to provide high quality information to customers using a natural language interface. Particularly, banking chatbots need to provide financial information to customers in rich detail. Daft and Lengel (1986) noted that information richness is a quality of information indicating its communication properties to the audience. Further, Balasubramaniam et al. (2003) showed that having high quality information on business websites has a positive effect on customer satisfaction. As chatbot gender forms part of the information being presented to a customer, it is meaningful to observe the effect of chatbot gender on perception of information quality or richness. General differences between genders have been discerned in terms of communication styles and ability. Women tend to be more expressive and know how to use nonverbal cues for clearer communication (Dennis et al., 1999). Previous studies have demonstrated that compared to those of men's, Women's communications are usually more detailed. Women tend to use more words, and their expressions are more descriptive and emotional (Niedźwieńska, 2003). Due to the detailed characteristics of women communications, it is noted that females produce more reliable accounts in terms of information richness than men's (Nahari and Pazuelo, 2015). These notable differences in terms of gender can also be found in digital forms of communication (Chou and Tsai, 2007). Thus, we propose;

Hypothesis 4: A customer encountering a chatbot with a female name (compared to a male name) will report higher perceived Information Richness (IR) in a chatbot service interaction.

2.2.5 Chatbot Gender and Intimacy

Intimacy indicates a set of interactions in which people engage in order to feel understood and cared for. The resulting attachment to others is brought on partially by the disclosure of information (Widener, 2019). When customers are in intimate relationships with a business, they feel a strong affective connection (Bügel et al., 2010). Hansen et al. (2003) find that high consumer-business intimacy positively correlates with consumers' willingness to share information with service providers, and that banking chatbots are dealing with customers' personal information. Chatbots enjoy two distinct advantages in intimacy over websites. Firstly, chatbots are installed on personal messenger apps that are much easier and more convenient to access and communicate on than a website, hence, chatbot occupies a very "intimate and trustful space" (Boutin, 2017). Secondly, since an AI chatbot uses natural language processing to converse with users in a more personal and human way, this leads to a more intimate, personalized experience than navigating a website. One advantage of chatbots lies in their ability to achieve intimacy with customers, something female chatbots tend to be better at accomplishing. It is perceived that women express their feelings better and enjoy intimate conversations more than men do, as their conversational styles are oriented towards developing intimacy in order to maintain relationships. More specifically, women tend to make more efforts to facilitate conversations. They try harder to agree and acknowledge points made by others through further talk and asking questions (Dennis et al. 1999). Based on this understanding of the tendency of women towards a more intimate communication style and its applications to gender assignment decisions when creating banking chatbots, we propose;

Hypothesis 5: A customer encountering a chatbot with a female name (compared to a male name) will report higher perceived Intimacy (IC) in a chatbot service interaction.

This study also investigates the effect of the gender congruence between chatbot and customer on customer responses. Similarity-attraction effect is the widespread tendency of people to be attracted to others who are similar to themselves in important respects (Reis, 2007). A common explanation for similarity-attraction effect is that interpersonal similarity leads to validation of one's personal traits and opinions via consensus support (Byrne and Clore, 1970). Research points to the fact that when users discover similarities in demographics (including gender), interests, and attitudes, they become more attracted to each other (Montoya et al., 2008). Although Grohmann (2009) found that congruence between brand gender and consumers' sex identity has a positive effect on consumer responses to brands, this is not always true. Söderberg (2013) found both male and female consumers were more likely to follow advice given by a female financial advisor, implying that gender congruence was not having a positive effect. As research is lacking consensus on the effect of gender congruence in the chatbot medium, we are thus interested in the effect of the congruence of chatbot gender and respondent gender on the five measures described above: Customer Satisfaction (CS), overall Brand Evaluation (BE), Brand Conceptual Fluency (BCF), perceived Information Richness (IR),and Intimacy (IC). We define congruent as the scenario where both chatbot and respondent genders are the same. Hence, we propose;

Hypothesis 6: A customer exposed to a congruent gender chatbot will report higher Customer Satisfaction(CS) in a chatbot service interaction.

Hypothesis 7: A customer exposed to a congruent gender chatbot will report higher overall Brand Evaluation(BE) in a chatbot service interaction.

Hypothesis 8: A customer exposed to a congruent gender chatbot will report *higher Brand Conceptual Fluency(BCF) in a chatbot service interaction.*

Hypothesis 9: A customer exposed to a congruent gender chatbot will report higher Information Richness(IR) in a chatbot service interaction.

Hypothesis 10: A customer exposed to a congruent gender chatbot will report higher Intimacy(IC) in a chatbot service interaction.

2.3 Brand Communication Style

Choosing and crafting the appropriate communication style for different channels is an important part of any brand's marketing strategy. Gretry et al. (2017) found consumers perceived an informal style of communication as inappropriate for unfamiliar brands. Using a communication style on one channel that conflicts with overall brand communication style can lead to confused customers and higher drop-off rates. As communication style is an important personality trait, different communication styles can thus be examined under the lens of brand personality.

The five brand personality dimensions proposed by Aaker (1997) include sincerity, excitement, competence, sophistication and ruggedness, each with its own subset of categories. As the dimensions are likely to be associated with human characteristics, they are helpful to differentiate markets and develop relationships with consumers (Aaker, 1997). However, Aaker's five brand personality dimensions do not include other human characteristics such as age and gender although she covers those concepts in loose definition of brand personality (Azoulay & Kapferer, 2003). Furthermore, the five factors could not be applied consistently across cultures. To resolve some of these issues, Geuens et al. (2009) developed a new measure for brand personality incorporating twelve items and five factors (Activity, Responsibility, Aggressiveness, Simplicity, and Emotionality).

For the purpose of our study though, we are not concerned so much with which brand personality dimension scale is more accurate, but rather the specific personality's communication styles that banks would likely use to differentiate their chatbots. In particular, this study is focused on the “responsibility”, and “activity”, communication styles. These styles are more likely to impact high credence services - services like banking that require higher degree of knowledge and training to execute (Ostrom and Iacobucci, 1995) as well as high degree of proactive client relationship management (Gordon et al., 2016).

Ⅲ. Methodology

3.1 Experiment Design and Instrument

This study adopted a 2 by 2 between subject factorial design manipulating the chatbot gender (male vs female) and Congruence of Chatbot Gender and Respondent Gender (CCGRG) (congruent vs incongruent). A scenario-based experiment was conducted for two weeks in June 2017 with study participants residing in the United States recruited from Amazon M-Turk. The participants were randomly assigned to one of four scenarios. The experimental stimuli first introduced the scenario of a fictitious bank, ACE Bank, which had recently launched an AI Chatbo. Study participants were then provided with an instruction to imagine themselves as customers of ACE Bank interacting with the new AI Chatbot Chip(Male) / Grace(Female) and view a chatbot interaction script. The chatbot replies were edited to include words and phrases that convey a Responsible/ Active communication style. All other information in the scripts across four conditions including savings amount, expenditures, budgeting, and customer queries were kept constant. Two names, Chip and Grace, were selected as they are commonly associated with male and female genders. After reading introduction of the survey, the brand manifesto, and viewing a randomized picture of the chatbot scenario, participants responded to a series of questions. To increase the reality of the script stimuli, we based the script scenario on a real financial chatbot's (Cleo) conversation simulation where a customer queried their account balance, budget, and expenditure for the month(https://meetcleo.com/). We also substituted certain keywords as advised by native English speakers to fit the local cultural context and enhance realism. Chatbot windows and chat bubbles are carefully designed to mimic those of Facebook Messenger as close as possible.

3.2 Participants

One hundred thirty respondents participated in the study and their demographic information is provided in Table 1. Seventy one respondents were males (54.6%) while the age range was between 24 and 69 (M=43.93, Standard Deviation (SD)=10.58). For the question "How many times have you used Chatbot Financial services in the past 6 months?", the frequency of using online customer chat service was Mchatbot frequency =0.97.

<Table 1> Participant Characteristics (n = 130)

	n	%
Education		
2 year degree	11	8.46
4 year degree	48	36.92
Doctoral Degree	4	3.08
High School / GED	11	8.46
Master's Degree	27	20.77
Professional Degree	2	1.54
Some College	27	20.77
Race		
White/Caucasian	101	77.69
African Americans	8	6.15
Hispanic	7	5.38
Asian	7	5.38
Pacific Islander	2	1.54
Others	5	3.85

3.3. Measurements

Customers' overall perception of ACE Bank's brand after viewing the stimuli was measure by a four-item scale of Customer Satisfaction (CS). We adopted a four-item scale from Dawar and Pillutla (2000) to measure overall Brand Evaluation (BE), a three-item scale from Sirianni et al. (2013) to measure Brand Conceptual Fluency (BCF), a six-item scale adapted from Patrakosol and Lee (2013) to measure perceived information richness (IR), and a five-item scale adapted from Turel et al. (2013) to measure perceived Intimacy (IC). Table 2 summarizes the items comprising each construct. All questions were measured on a 7-point Likert scale (1=strongly disagree and 7=strongly agree).

The measurement constructs' respective Cronbach alpha values, Average

Variance Extracted (AVE) values, and Composite Reliability values are listed in Table 3. All the constructs' Cronbach alpha values are above 0.7, implying they have good reliability. Similarly, all AVE values are above 0.5, and all composite reliability values are above 0.7, implying our constructs have good validity. Participants perceived the scenarios realistic (M=5.48, SD=1.506). The data collected from the questionnaires were analyzed using SPSS by performing an analysis of variance (ANOVA).

<Table 2> Construct Measurement Items

Construct	Item	Original Source
Customer Satisfaction (CS)	As the customer in the Chatbot interaction scenario, what is your overall impression of the service experience? Negative… Positive Unappealing… Appealing Bad… Good As the customer in the Chatbot interaction scenario, how would you evaluate this service experience? Unacceptable…Outstanding	
Overall Brand Evaluation (BE)	Based on the Chatbot interaction scenario you read, how do you feel about ACE Bank's brand overall? 1. Dislike…Like 2. Not at all trustworthy…Very trustworthy 3. Very low quality…Very high quality 4. Not at all desirable…Very desirable	(Dawar & Pillutla, 2000)
Brand Conceptual Fluency (BCF)	1. I have a clear understanding of what this bank's brand stands for. 2. It was easy for me to identify what this bank's brand represents to customers. 3. It was easy for me to describe what this bank's brand represents to customers.	(Sirianni et al., 2013)
Information Richness (IR)	1. Overall, the information provided by the Chatbot was effective. 2. The Chatbot provides accurate information. 3. I find the Chatbot provides easy to understand information. 4. The Chatbot provides information at the right level of detail. 5. The Chatbot provides information in a confusing way. 6. The Chatbot conversation communicated information about the values, beliefs and culture of the bank.	(Patrakosol & Lee, 2013)
Intimacy (IC)	1. The Chatbot provided the service in a friendly manner. 2. The Chatbot appeared enthusiastic about helping me. 3. The Chatbot seems to care about me.	(Turel et al., 2013)

	4. The Chatbot seems to be self-centered. 5. The Chatbot seems to be understanding.	

<Table 3> Constructs' Cronbach Alpha values and AVE value

Construct	Cronbach's Alpha	AVE	Composite Reliability	Mean	SD
CS	0.941	0.851	0.958	5.9942	1.11389
BE	0.933	0.832	0.952	5.6827	1.10355
BCF	0.925	0.87	0.953	5.2692	1.13294
IR	0.775	0.506	0.85	5.5115	0.75631
IC	0.794	0.575	0.865	5.5846	0.82916

Ⅳ. Results

4.1 The Effect of Chatbot Gender and the Congruence of Chatbot Gender and Respondent Gender on Customer Responses

To test the effect of chatbot gender and the congruence of chatbot gender and respondent gender on various customer responses, a series of ANOVAs was performed with 2 (Chatbot Gender type: Male vs Female) by 2 (CCGRG type: Congruent vs Incongruent) between-subject factors including five customer response constructs. Table 4 presents the summary of ANOVA results of the effect of chatbot gender and CCGRG on five dependent variables; customer responses. First, the main effect of chatbot gender has a significant effect on customer satisfaction (CS) ($F=5.9$, $p<0.05$), and brand conceptual fluency (BCF) ($F=10.913$, $p<0.05$). It indicates that the effect of the certain chatbot gender compared to the other gender type exist for enhancing customer satisfaction and brand conceptual fluency. Based on the related theories, this study hypothesized that CS and BCF would be higher when customers encounter the female chatbot. However, participants who were exposed to the male chatbot, Chip, reported significantly higher mean scores than those exposed to the female chatbot, Grace on CS (Chip: $M=6.24$, $SD = 0.11$/ Grace: $M=5.75$, $SD = 0.16$) and BCF (Chip: $M=5.60$, $SD = 0.13$/ Grace: $M=4.95$, $SD = 0.14$). Table 5 shows the mean responses to two chatbots types respectively. Hypotheses 1 and 3 are thus not supported because the results are significant but opposite to what was expected. In addition, as Table 4 indicates, chatbot gender has no significant effect on BE, IR, and IC, hence, Hypotheses 2, 4, and 5 are not supported.

<Table 4> F values for ANOVA of CS, BE, BCF, IR and IC with Chatbot Gender and Congruence of Chatbot Gender and Respondent Gender (CCGRG)

	CS	BE	BCF	IR	IC
Chatbot Gender	5.9*	2.757	10.913*	1.251	0.33
CCGRG	0.147	0.244	0.003	0.523	0.072
Chatbot Gender*CCGRG	1.232	0.656	0.006	5.67*	8.947*

*$p<0.05$

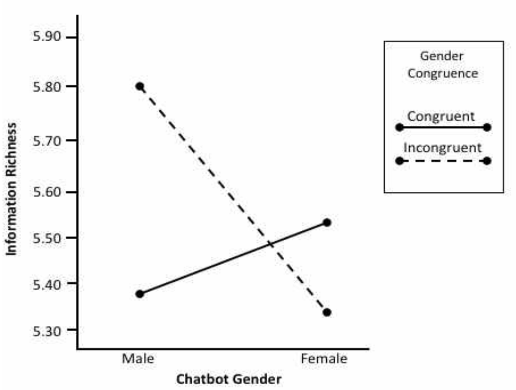
<Table 5> Mean Responses to Chatbot Gender Type

Chatbot Gender	CS	BE	BCF	IR	IC
Male	6.24	5.86	5.60	5.59	5.64
	(0.11)	(0.12)	(0.13)	(0.09)	(0.10)
Female	5.75	5.51	4.95	5.43	5.53
	(0.16)	(0.15)	(0.14)	(0.09)	(0.10)

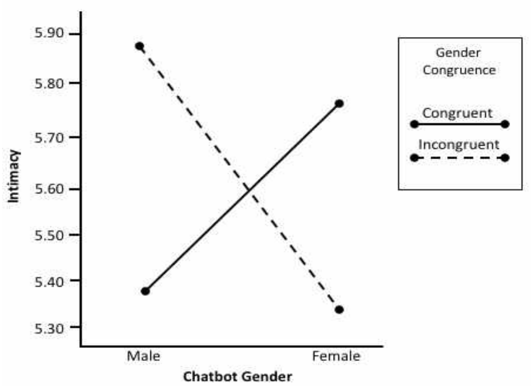
Standard errors are in parentheses.

This study assumed that when a customer encounters a chatbot with the same gender would have higher CS, BE, BCF, IR, and IC. However, as it is shown on Table 4, CCGRG has no significant effect on any of the response measures, which means encountering a chatbot with the same gender did not matter for participants' reponses. Therefore, the results do not support Hypotheses 6, 7, 8, 9, and 10. Although there was no significant effect found for CCGRG on any response measures, it is noted that there was a significant interaction effect between chatbot gender and CCGRG on information richness (IR) ($F=5.67$, $p<0.05$); and perceived intimacy (IC) ($F=8.947$, $p<0.05$) (see Table 4). This indicates that the male chatbot demonstrates significantly higher IR ($F=5.035$, $p<0.05$), and IC ($F=5.433$, $p<0.05$) scores when viewed by a gender incongruent (female) participant. In contrast, a female chatbot does not show significantly different IR ($F=1.32$, $p>0.05$) and IC ($F=3.63$, $p>0.05$) scores when viewed by a gender congruent (female) participant as compared to a gender incongruent (male) participant (see Figures 1-2).

<Figure 1> Interaction effect of Chatbot Gender*Congruence of Chatbot Gender and Respondent Gender on Information Richness (IR).



<Figure 2> Interaction effect of Chatbot Gender*Congruence of Chatbot Gender and Respondent Gender on perceived Intimacy (IC).



4.2 Chatbot Gender; Chatbot Communication Style

To test the effect of the chatbot genders, and responsibility and activity communication styles, another set of ANOVAs was performed. Table 6 summarizes the ANOVA results of a 2 (Chatbot Communication Style type: Responsibility vs. Activity) x 2 Chatbot Gender: male vs. female) design construct. Results show that there is no significant main effect of chatbot communication style and interaction effect between chatbot gender and chatbot communication style on any of the measures. This indicates that chatbot's communication style either responsible or activity did not influence customers responses such as customer satisfaction and brand evaluation. Also, the impact of chatbot's communication style on customer responses was not different based on the chatbot gender.

<Table 6> F values for ANOVA of CS, BE, BCF, IR and IC with Chatbot Gender and Chatbot Communication Style

	CS	BE	BCF	IR	IC
Chatbot Gender	6.405*	3.045	10.956*	1.24	0.451
Chatbot Communication Style	0.003	0.023	0.004	1.24	0.118
Chatbot Gender*Chatbot Communication Style	0.004	0.093	0.047	1.886	3.855

*$p<0.05$

V. Discussion and Conclusion

5.1. Theoretical and Practical Implications

The concept of gender in the customer service context has been examined with regards to the different characteristics and stereotypical roles customers associate with the two genders (Niculescu et al., 2010; Pradhan et al, 2017). This study has looked at the effect of assigning particular genders to chatbots, as well as the effect of gender congruence between banking chatbots and bank customers. Based on the related literature, we assumed that adopting female chatbots would enhance customer responses such as satisfaction and brand evaluation. However, our study found that a chatbot with the male name significantly increases customer satisfaction and brand conceptual fluency. Customer satisfaction is the result of diverse elements, and this study assumed that friendly and polite female chatbots who are more relationship oriented by comparison to male chatbots might have an advantage when it comes to achieving customer satisfaction (Verhagen et al., 2014). With results that point in the opposite direction, however, it may be that customer satisfaction when it comes to banking services is more a result of the efficiency of service process and the comprehensiveness of information provided to the customer (Wolfinbarger and Gilly, 2003). As banks and other financial institutions exist for the purpose of managing assets, customers may prefer that their services be focused on functional qualities such as accuracy and speed.

Brand conceptual fluency indicates the ease with which a customer is able to understand and process business information (Sirianniet al., 2013). It is known that brand conceptual fluency can be improved when the target is primed or when information is presented in a predictive manner (Lee and Labroo, 2004). As female chatbots are being adopted more frequently for the banking context, this study assumed that customers perceive the role of financial assistant as a stereotypically female one. But, given the study's results, it is worth questioning whether this truly is a stereotype. Traditionally, the role of "banker" is most frequently occupied by males; this might explain the fact that male chatbots were better received. An alternative explanation is that the tasks the banking chatbots were called upon to execute may have been perceived differently by the participants of the study.

This is an interesting preliminary finding given that banks and fintech chatbots predominantly use female names. The advantages this study found regarding male chatbots' abilities to procure better customer responses were unexpected; this finding may be useful to banks that are both already utilizing and considering the adoption of chatbot services. Banking chatbots, which tend most frequently to be assigned female names, do not necessarily need to be represented as female. We suggest instead that banks should find names for their chatbots that are likable, creative, and unique regardless of gender.

This study found the significant interaction effect between the chatbot gender and Congruence of Chatbot Gender and Respondent Gender (CCGRG) on information richness and intimacy. Based on the similarity-attraction effect—a tendency to be attracted to people with the same or similar major perspectives (Reis, 2007) —this study predicted that customer responses would be improved when customers encounter a chatbot of the same gender. However, the results show that the female chatbot had higher information richness and intimacy scores when interacted with a male participant (incongruent gender). This finding is inconsistent with similarity-attraction effect, and it is interesting that this effect does not hold for male chatbots. Although the results were not significant, a female chatbot also had higher information richness and intimacy scores when viewed by a gender congruent (female) participant. The results might provide a signal that the communications of female chatbots are perceived more personal and detailed (Dennies et al., 1999; Nahari and Pazuelo, 2015). Overall, this finding suggests banks interested in improving perceived intimacy and information richness of their chatbots with clients should first identify the clients' gender and assign the appropriate chatbot gender accordingly. Specifically, a bank targeting male clients might want to assign a female chatbot to maximize perceived intimacy and information richness. In contrast, using a male chatbot would have no significant difference in perceived intimacy and information richness between male and female clients.

The effect of chatbot gender on customer responses did not differ between chatbot communication styles (activity communication style vs. responsibility communication style). The interaction between chatbot communication style and chatbot gender also had no significant effect on the measures. This could be because our scripted stimuli are not sufficiently differentiated to cause significant differences in customer responses and thus still needs further investigation in the future. However, with this result, it is suggested that it might not be effective or even feasible for banks to differentiate their chatbots with different communication styles. It is time-consuming and difficult to craft a unique communication style using specific keywords and phrases that actually are perceived differently by customers in a chatbot environment. Our results suggest focusing on the dimension of gender in a chatbot could be a simpler and more effective way for banks to differentiate their chatbots and improve perceived intimacy, information richness, customer satisfaction, and brand conceptual fluency.

Banks, like other service-oriented businesses, are interested in improving customer satisfaction levels, brand awareness, and brand perceptions. It is found that financial institutions that had inadequate leadership support for the brand were poorly differentiated, exhibited a lack of understanding and confusion about branding issues, and had service quality problems (de Chernatony and Cottam, 2006). Therefore banks should have broad interest in strengthening their brands and leveraging them in their customer service interactions including on their chatbot platforms. Tailoring a suitable gender on their chatbot can thus help a bank differentiate its brand and improve service quality.

5.2 Limitations and Future Study

This study has some limitations that can be improved for future research. First, our stimuli used for the experiment are static image simulations of Facebook Messenger chat windows which are not as authentic as a real chatbot. It would be ideal if future research could build a chatbot type used in the field and test it so that participants have a more authentic stimuli simulating a real-world chatbot interaction. The likeability of “Chip” and “Grace” as male and female names can be a confounding variable. A separate pilot study is needed to measure likeability of different male and female names. This would allow us to use a male and female name of equivalent likeability in the main study. Lastly, we adopted measures such as customer satisfaction, overall brand evaluation, brand conceptual fluency, perceived intimacy, and information richness to describe customer responses in our study design. While this can explain aspects of branding research, marketers are also exploring how chatbots can influence purchase decisions (Matthews, 2017). Most existing chatbots already offer some forms of choice in presenting product options and decision options to users. We foresee several future research opportunities investigating how chatbot gender and customer gender impacts customer decision-making in a chatbot environment.

References

Aaker, J. L. (1997), Dimensions of Brand Personality. *Journal of Marketing Research, 34*(3), 347−356.

Alharbi, S. (2020). How Does Contact With an Artificial Intelligence Avatar Influence Customer Perceptions of Bank Service Quality? (Doctoral dissertation, Auckland University of Technology).

Azoulay, A. and Kapferer, J. (2003). Do brand personality scales really measure brand personality? *Journal of Brand Management, 11*(2), 143−155.

Balasubramanian, S., Konana, P. and Menon, N. M. (2003), Customer Satisfaction in Virtual Environments: A Study of Online Investing. *Management Science, 49*(7),871−889.

Bapat, D., & Thanigan, J. (2016). Exploring relationship among brand experience dimensions, brand evaluation and brand loyalty. *Global Business Review, 17*(6), 1357−1372.

Boutin, P. (2017, April 4), Does a Bot Need Natural Language Processing? - Chatbots Magazine. Retrieved from https://chatbotsmagazine.com/does−a−bot−need−natural−language−processing−c2f76ab7ef11

Brakus, J.J., Schmitt, B.H., & Zarantonello, L. (2009). Brand experience: What is it? How is it measured? Does it affect loyalty? *Journal of Marketing, 73*(3), 52-68.

Byrne, D. and Clore, G. L. (1970), Are inforcement model of evaluative responses. *Personality: An International Journal, 1*, 103−128.

Byun, S.−H. & Cho, C.−H. (2020). The Effect of the Anthropomorphism Level and Personalization Level on AI Financial Chatbot Recommendation Messages on Customer Response. *The Korean Journal of Advertising and Public Relations, 22*(2), 466−502.

Bügel, M. S., Buunk, A. P. and Verhoef, P. C. (2010). A Comparison of Customer Commitment in Five Sectors Using the Psychological Investment Model. *Journal of Relationship Marketing, 9*(1), 2−29.

Chou, C., & Tsai, M. (2007). Gender differences in Taiwan high school students' computer game playing. *Computers in Human Behavior, 23*(1), 812-824.

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2018). Chatbot e−service and customer satisfaction regarding luxury brands. *Journal of Business Research. 117*, 587-595.

Cleo AI Ltd. (2017), Cleo, an intelligent assistant for your money. Retrieved from https://meetcleo.com/

Daft, R. L. and Lengel, R. H.(1986), Organizational Information Requirements, Media Richness and Structural Design. *Management Science, 32*(5), 554-571.

Dale, R. (2016), The return of the chatbots. *Natural Language Engineering, 22*(5),811-817.

Dawar, N. and Pillutla, M. M.(2000), Impact of Product-Harm Crises on Brand Equity: The Moderating Role of Consumer Expectations. *Journal of Marketing Research, 37*(2),215-226.

De Chernatony, L. and Cottam, S. (2006), Why are all financial services brands not great? *Journal of Product & Brand Management, 15*(2), 88-97.

Dennis, A. R., Kinney, S. T., & Hung, Y. T. C. (1999). Gender differences in the effects of media richness. *Small Group Research, 30*(4), 405-437.

Forlizzi, J., Zimmerman, J., Mancuso, V., & Kwak, S. (2007, August). How interface agents affect interaction between humans and computers. In Proceedings of the 2007 conference on Designing pleasurable products and interfaces (pp. 209-221).

Geuens, M., Weijters, B. and DeWulf, K. (2009), A new measure of brand personality. *International Journal of Research in Marketing, 26*(2), 97-107.

Gordon, R., Zainuddin, N. and Magee, C. (2016), Unlocking the potential of branding in social marketing services: utilising brand personality and brand personality appeal. *Journal of Services Marketing, 30*(1), 48-62.

Gretry, A., Horváth, C., Belei, N., and Van Riel, A. C. (2017), "Don't pretend to be my friend!" When an informal brand communication style backfires on social media. *Journal of Business Research, 74*, 77-89.

Grohmann, B. (2009), Gender Dimensions of Brand Personality. *Journal of Marketing Research, 46*(1),105-119.

Gupta, S., Borkar, D., De Mello, C. and Patil, S. (2015), An E-Commerce Website based Chatbot. *International Journal Of Computer Science And Information Technologies, 6*(2), 1483-1485.

Hansen, H., Sandvik, K., and Selnes, F. (2003), Direct and Indirect Effects of Commitment to a Service Employee on the Intention to Stay. *Journal of Service Research, 5*(4),356-368.

Kim, J. W., Cho, H. I., & Lee, B. G. (2019). The Study on the Factors Influencing on the Behavioral Intention of Chatbot Service for the Financial Sector: Focusing on the UTAUT Model. *Journal of Digital Contents Society, 20*(1), 41-50.

Kim, B.-S. & Kim. J. (2020). Characteristics of customer networks that affects to satisfaction-loyalty relationships. *Journal of Korea Service Management Society, 21*(3), 149-176.

Kunze, L. (2016, February 2). On chatbots. Retrieved from https://techcrunch.com/2016/02/16/on-chatbots/.

Lee, A. Y., & Labroo, A. A. (2004). The effect of conceptual and perceptual fluency on brand evaluation. *Journal of Marketing Research, 41*(2), 151-165.

Lee, M. S. (2018). The effect of perceived calue on customer satisfaction and loyalty in payment service market : focusing on moderating effects of switching costs. *Journal of Korea Service Management Society, 19*(4), 47-71.

Leong, L. Y., Hew, T. S., Lee, V. H., & Ooi, K. B. (2015). An SEM-artificial-neural-network analysis of the relationships between SERVPERF, customer satisfaction and loyalty among low-cost and full-service airline. *Expert Systems with Applications, 42*(19), 6620-6634.

Lieven, T., Grohmann, B., Herrmann, A., Landwehr, J. R. and Van Tilburg, M. (2014), The Effect of Brand Gender on Brand Equity. *Psychology & Marketing, 31*(5),371-385.

Lin, X., Featherman, M., Brooks, S. L., & Hajli, N. (2019). Exploring gender differences in online consumer purchase decision making: An online product presentation perspective. *Information Systems Frontiers, 21*(5), 1187-1201.

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019), Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science, 38*(6), 937-947.

Matthews, K. (2017, May 12), 3 Reasons Why 2017Is The Year Of The Chatbot. Retrieved from https://chatbotsmagazine.com/3-reasons-2017-is-the-year-of-the-chatbot-6fa0c783a444

McDonnell, M., & Baxter, D. (2019). Chatbots and gender stereotyping. *Interacting with Computers, 31*(2), 116-121.

Montoya, R. M., Horton, R. S. and Kirchner, J. (2008), Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships, 25*(6), 889-922.

Morris, R. R., Kouddous, K., Kshirsagar, R., & Schueller, S. M. (2018). Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. *Journal of medical Internet research, 20*(6), e10148.

Nagarhalli, T. P., Vaze, V., & Rana, N. K. (2020, March). A Review of Current Trends in the Development of Chatbot Systems. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 706-710). IEEE.

Nahari, G., & Pazuelo, M. (2015). Telling a convincing story: Richness in detail as a function of gender and information. *Journal of Applied Research in Memory and Cognition, 4*(4), 363-367.

Niculescu, A., Hofs, D., Van Dijk, B., & Nijholt, A. (2010, December). How the agent's gender influence users' evaluation of a QA system. In 2010 International Conference on User Science and Engineering (i-USEr) (pp. 16-20). IEEE.

Niedźwieńska, A. (2003). Gender differences in vivid memories. *Sex Roles, 49*(7-8), 321-331.

Okoro, A. S. (2014). Impact of electronic banking instruments on the intermediation efficiency of the Nigerian economy. *International Journal of Accounting Research, 1*(6), 14-24.

Ostrom, A. and Iacobucci, D. (1995), Consumer Trade-Offs and the Evaluation of Services. *Journal of Marketing, 59*(1), 17-28.

Park, D.-A. (2017), A Study on Conversational Public Administration Service of the Chatbot Based on Artificial Intelligence. *Journal of Korea Multimedia*

Society, 20(8), 1347-1356.

Patrakosol, B. and Lee, S. M.(2013), Information richness on service business websites. *Service Business, 7*(2), 329-346.

Pavlus, J. (2016, May 1), The Next Phase of UX: Designing Chatbot Personalities. Retrieved from https://www.fastcodesign.com/3054934/the-next-phase-of-ux-designing-chatbot-personalities

Pradhan, D., Kapoor, V., & Moharana, T. R. (2017). One step deeper: gender and congruity in celebrity endorsement. *Marketing Intelligence & Planning, 35*(6), 774-788.

Ravald, A. and Grönroos, C. (1996), The value concept and relationship marketing. *European Journal of Marketing, 30*(2), 19-30.

Reis, H. T. (2007), Similarity-Attraction Effect. In Encyclopedia of Social Psychology. Retrieved from http://dx.doi.org/10.4135/9781412956253.n517

Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., & Mazurek, G. (2019). In bot we trust: A new methodology of chatbot performance measures. *Business Horizons, 62*(6), 785-797.

Ranoliya, B. R., Raghuwanshi, N., & Singh, S. (2017, September). Chatbot for university related FAQs. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1525-1530). IEEE.

Sirianni, N. J., Bitner, M. J., Brown, S. W. and Mandel, N. (2013). Branded service encounters: strategically aligning employee behavior with the brand positioning. *Journal of Marketing, 77*(6), 108-123.

Söderberg, I. (2013), Relationships between advisor characteristics and consumer perceptions. *International Journal of Bank Marketing, 31*(3), 147-166.

Song, Y. J. & Choi, S. M. (2020). The effects of chatbots' anthropomorphism and self-disclosure on consumers' perceptions of and attitude toward the chatbots. *Journal of the HCI Society of Korea, 15*(1), 17-28.

Suh, C. J. & Yoon, J. O. (2019). The effects of perceived chatbot service quality on customer satisfaction and word of mouth. *Journal of Korea Service Management Society, 20*(1), 201-222.

Torba, A. (2016, May 24), Why Chatbots Are Exactly What Digital Marketers

Need Right Now. Retrieved from https://chatbotsmagazine.com/why-chatbots-are-exactly-what-digital-marketers-need-right-now-bcdc339d56f6#.olasoaa48

Turel, O., Connelly, C. E. and Fisk, G. M. (2013), Service with an e-smile: Employee authenticity and customer use of web-based support services. *Information & Management, 50*(2-3), 98-104.

Verhagen, T., Van Nes, J., Feldberg, F., & Van Dolen, W. (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication, 19*(3), 529-545.

Wang, L. C., Baker, J., Wagner, J. A. and Wakefield, K. (2007), Can a retail web site be social? *Journal of Marketing, 71*(3), 143-157.

Weizenbaum, J. (1966), ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36-45.

Widener, A. (2019). Need to belong, privacy concerns and self-disclosure in chatbot Artificial Intelligence interaction. (Master dissertation, Ewha Woman's University).

Wolfinbarger, M. F. & Gilly, M. (2003), "eTailQ: dimensionalizing, measuring, and predicting e-Tail quality", *Journal of Retailing, 79*(3), pp. 183-198.

* 저자소개 *

· 산지드 비케이(sanjeed27@gmail.com)

싱가포르의 빠르게 성장하고 있는 구강케어 전문 스타트업 업체인 Zenyum의 Head of Content이자 SEO이다. 대만 국립 칭화대학교에서 MBA학위를 취득하였다.

· 김 민 경(mkim@hongik.ac.kr)

현재 홍익대학교 상경대학 상경학부 조교수로 재직 중이며, 주요 강의와 연구 분야는 소비자행동에 기반을 둔 서비스와 하스피탈리티 마케팅이다. Pennsylvania State

University에서 하스피탈리티 경영 박사학위를 취득하였다.

· Chenya Wang(cywang@iss.nthu.edu.tw)

현재 대만 국립 칭화대학교의 서비스 사이언스 대학의 부교수로 재직 중이며, Pennsylvania State University에서 하스피탈리티 경영 박사학위를 취득하였다. 주요 강의 및 연구 분야는 서비스마케팅, cross-cultural differences, 그리고 서비스 커뮤니케이션에서의 AI 관련 주제이다.