THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學
Pao Yue-kong Library
包玉剛圖書館

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

Pao Yue-kong Library, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

http://www.lib.polyu.edu.hk

TWO STUDIES OF HUMAN-AI INTERACTIONS

WEN XIE

PhD

The Hong Kong Polytechnic University

2023

The Hong Kong Polytechnic University

Department of Management and Marketing

Two Studies of Human-AI Interactions

Wen Xie

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

July 2023

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma except where due acknowledgment has been made in the text.

Signature:

Name: Xie Wen

Abstract

In light of the vital role of artificial intelligence (AI) and the distinctive characteristics of AI (such as anthropomorphism, inexplicability, and natural language), AI can be extremely beneficial to us. However, human-AI interaction inefficiency is very common (Agomuoh, 2023; Cem, 2023). It is necessary for us to better understand how users and AI communicate and how users feel in a world where humans and AI co-exist.

This thesis consists of two AI research. This first study discusses how chatbot designs affect the human-chatbot interaction process and, thus, influence human-AI interaction outcomes. Chatbots are gaining momentum in a variety of business functions, such as IT service, human resource management, customer service, and sales. Although chatbot proactivity design and social identity design are popular in the industry, limited existing research on chatbots has investigated chatbot proactivity design and its role as a boundary condition of chatbot social identity design. In order to optimize user evaluation on the service with chatbots, there is thus a need from both the business and the academia aspects to better understand how chatbot social designs affect the human-chatbot conversational process and, in turn, affect user perception and task success from the perspective of uncertainty reduction. The field experiment results reveal that proactivity design will decrease communication inefficiency and, thus, increase customer satisfaction and the probability of task success. Furthermore, chatbot social identity design weakens the negative effect on customer satisfaction made by ineffective communication.

The second study investigates how the introduction of a hybrid human-AI service affects user evaluations of the service. AI and humans do have their own advantages when facing different tasks in digital platforms—for example, AI agents are better skilled in reliability, scalability, speed, accuracy, and generalization, while human agents are better skilled in creativity, judgment, and empathy (Rai and Sarker, 2019). A Human-AI hybrid service system can benefit from both AI and Human agents' advantages, and it is of interest to study the impact of human-AI hybrid services in digital platforms. In particular, our study examines the associations between human-AI hybrid services and user evaluations in the context of a major audio streaming platform. On the platform, the introduction of AI podcasters creates the phenomenon—for the same audio, sometimes the AI voice version is released first, and sometimes the human voice version is released first. This provides us with a great opportunity to study how the introduction of a human-AI hybrid service affects user evaluations of the service from a temporal perspective. The empirical findings show that the user evaluations of those who are first exposed to a human service are improved with the presence of the human-AI hybrid service, and the user evaluations of those who are first exposed to an AI service get worse with the presence of the human-AI hybrid service. Our findings have important implications for both digital platforms and AI developers.

Keywords: ***human-AI interactions, communication inefficiency, uncertainty reduction theory, human-AI hybrid service, digital platform, the contrast effect, user evaluations on the service***

Acknowledgements

I would like to take this chance to express my sincere appreciation to all the people who assisted me in going through My Ph.D. journey. Especially, I would like to express my gratitude to my supervisor Prof. Xin Xu. He has given me a lot of guidance in my research. I would also like to thank Prof. Eric W.T. Ngai, Dr. Yong Jin, and Dr. Yue Feng for the help and guidance given to me. Last but not least, I would like to thank my family, my parents, my husband, and my kid for the unconditional support they have given me.

Contents

List of Tables

List of Figures

Chapter 1 : Introduction

1.1. Research Motivation

Artificial Intelligence (AI) was first introduced in1950s. It refers to a system that learns from its surroundings and takes actions to increase its chances of achieving its objectives (Poole et al., 1998). Data storage and data richness have been greatly improved with the development of Cloud Computing and Internet of Things technologies, which further facilitated the development of AI technology in the past ten years. AI is generally used for reasoning, problem-solving, planning, knowledge representation, learning, natural language processing, and perception. Now, AI is gaining momentum in a variety of business functions, such as detecting diseases in healthcare, AI-powered sales assistants, fraud prevention in marketing, and autonomous vehicles in manufacturing[1]. According to a report, AI software revenue is expected to reach 126 billion US dollars by 2025[2]. Despite its vast market value, most research on AI focuses more on its capabilities of AI and how it is applied in various fields (Huang and Rust, 2018; Jiang et al., 2017; Serey et al., 2021). The research on the AI user side is limited, especially user evaluations of the service. In addition, few studies have delved into the human-AI interaction process. Thus, it is an excellent opportunity to investigate AI from users' perspectives, such as the effects of AI designs on the human-AI interaction process and how users feel in a human-AI coexistence world.

The considerable market value and benefits of AI made AI research valuable. I believe that AI will be part of our lives, including healthcare, daily work, lifestyle, social media, etc. However, it is revealed that AI is pushed back from the user side (Luo et al., 2019). Therefore, we should understand how to get along with AI and investigate AI from the user's perspective. My research motivation is derived from AI's huge value and the current research state of AI. This thesis will study human-AI interaction from the perspectives of how the AI system designs influence the human-AI conversational process and how user evaluations of the service are affected in the process of introducing a human-AI hybrid service to digital platforms.

1.2. Existing Artificial Intelligence Research

To gain a clear picture of the current state of AI research, social science AI studies are reviewed. First, a word cloud is displayed to give us a basic understanding of AI research, such as most discussed algorithms and topics. Next, the literature review concluded three streams of AI research and identified the research gaps. Articles are selected from the Web of Science and Google scholar using the keyword—Artificial Intelligence. After ruling out technology paper (i.e., algorithm development), 48 social science studies are selected (i.e., Jarek and Mazurek, 2019; Jiang et al., 2017; Longoni et al., 2019). The word cloud is generated based on these studies (see Figure 1.1). We found that neural network, knowledge, algorithm, application, innovation, change, decision-making, and theory are the keywords in these paper. The word cloud figure reveals that neural network is the most discussed algorithm in social science AI paper. In addition, innovation, knowledge, changes, and AI applications are the most discussed AI topics in these AI studies in social science, which reveals existing AI research highly values what AI brings to us.

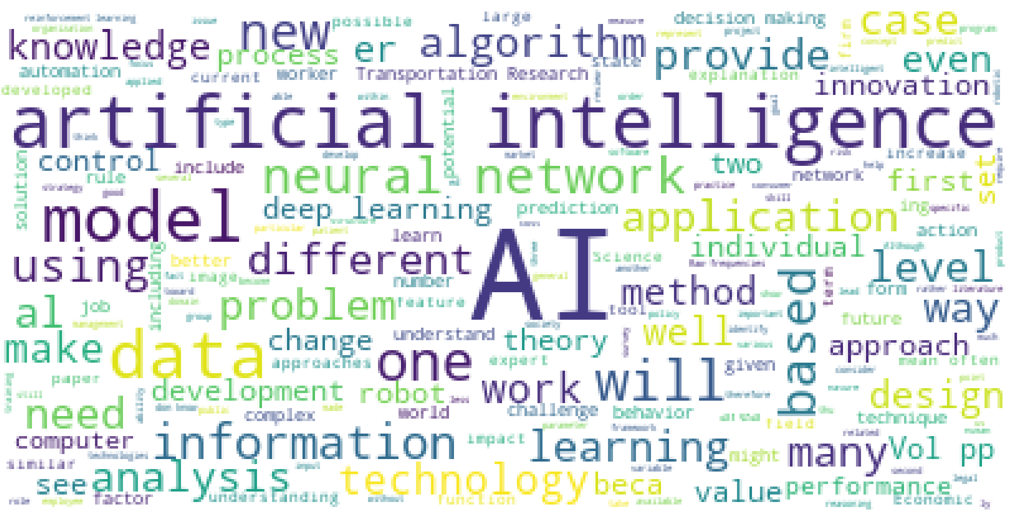

Figure 1.1. Word Cloud

Further, excepting lots of computer science studies, AI studies were mainly discussed in the following fields: marketing (e.g., Jarek and Mazurek, 2019; Wirth, 2018), healthcare (Brown et al., 2021; Jiang et al., 2017; Jin et al., 2020; Longoni et al., 2019), and management (Hussain et al., 2019; Raisch and Krakowski, 2021; Verganti et al., 2020; Webb et al., 2019).

After reviewing existing paper, there are three streams of AI research. One stream of research focuses on studying the consequences of AI adoption. Generally, there is no consensus on the impact of AI adoption, depending on the scenario in which it is used, such as the augmentation potential of AI adoption in employees' learning and development in the workplace (Wilkens, 2020), a positive effect of firm-level AI adoption on the firm outcome (Alekseeva et al., 2021), a negative effect of AI sales adoption on sales (Luo et al., 2019), no robust association between AI adoption among IT specialists and firm outcomes (Alekseeva et al., 2021), and consumers' low acceptance in medical AI adoption (Longoni et al., 2019). The second stream of research concentrates on discussing the capabilities of AI and possibilities for applying AI in a number of scenarios, such as AI in data management (Jiang et al., 2017; Jin et al., 2020; Serey et al., 2021), marketing segmentation (Raiter,2021), and AI models in healthcare (Pierce, 2019; Reddy et al., 2019). The last stream of AI research discusses the ethical problems of AI and, accordingly, how to build a safe AI by developing regulations, policies, or ethics (Calo, 2017; Fournier-Tombs, 2021; Yu et al., 2018). In this stream, researchers also revealed the phenomenon that people prefers human service provider to AI service provider (Longoni et al., 2019, Luo et al., 2019, Mou and Xu, 2017), which may result from the ethical problem of AI (Johnson and Verdicchio, 2017). Generally, AI research is still at an early stage. For example, in AI adoption consequences research, limited consequences are discussed. In addition, other research directions should also be taken into account. For instance, human-AI collaboration structure research is still at an early stage. The research on the human-AI interaction process is also limited. Although the self-disclosure of users is discussed in human-AI communication (Schanke et al., 2021), more human-AI communication processes (e.g., inefficiency) are worthy of being studied. Last but not least, existing studies put a lot of effort into AI technology development and AI applications (e.g., Jin et al., 2020; Prentice et al., 2020), while the research on the AI user side is limited. Therefore, this work discussed human-AI communication inefficiency and its impact on user satisfaction and task outcome. In the second study, we base on the common phenomenon of AI services and human services coexistence and conceptualize a human-AI collaboration construct—human-AI hybrid service. In this study, we investigate how the introduction of a human-AI hybrid service affects user evaluations.

1.3. Research Introduction

This thesis consists of two AI studies. Both studies investigate human-AI interactions from different perspectives. Given the AI literature research gap identified, this thesis investigates AI from the user perspective using different methods. The first study conducts an online field experiment and focuses on the human-AI interaction process, antecedents, and associated outcomes. The second study takes advantage of the human-AI hybrid service introduction to study the effects of human-AI hybrid service on user evaluations from a temporal perspective. Meanwhile, this thesis's purpose is to provide foundations for future human-AI interaction studies and call for attention to user-side research in AI research.

In the first study, chatbots' social designs are researched. Chatbots are the essential interfaces representing the companies to communicate with customers. Therefore, the social designs of chatbots are deserved to be investigated. In a famous beer company, we conducted an online field experiment on their social media platform. By encouraging participants to chat with our designed chatbots, we can collect and analyze their chats. This study applies uncertainty reduction theory to examine how chatbot social designs influence human-AI conversational processes and, thus, affect associated outcomes. This study focuses on two social chatbot designs—proactivity and social identity designs. Both two designs are popular in practice and have necessary practical implications. However, in academia, limited studies researched the two social designs. In addition, this study further investigates the human-AI interaction process—communication inefficiency, which is an under-researched but essential area in human-AI interaction research. This research will be displayed in Chapter 2.

The second study investigates how user evaluations of the service change in the process of introducing human-AI hybrid services to digital platforms. AI and humans do have their own advantages when facing different tasks in digital platforms—for example, AI agents are better skilled in reliability, scalability, speed, accuracy, and generalization, while human agents are better skilled in creativity, judgment, and empathy (Rai and Sarker, 2019). A human-AI hybrid service system can benefit from both AI and human agents' advantages, and it is of interest to study the impact of human-AI hybrid services in digital platforms. It is increasingly common to offer both AI and human services to users. In the context of an online streaming platform, we use the interesting phenomenon—AI agents and human agents provide the same service to users—to examine the effect of the human-AI hybrid service on user evaluations. For the same audio, users can listen to it broadcasted by either AI or humans. How do user evaluations of the service change when the human-AI hybrid service becomes available? This research examines the effect of the human-AI hybrid service introduction from a temporal perspective and reveals that the introduction of the human-AI hybrid service polarize user evaluations of the service. It sheds light on the underlying mechanism of the effects of human-AI hybrid service on user evaluations based on the contrast effect. This research will be displayed in Chapter 3.

In this thesis, I first introduce my research motivation and the current state of AI research in social science. Next, in chapter 1.3, the two research is briefly introduced. Chapters 2 and 3 will demonstrate the two studies. In the last chapter, the conclusions and future research will be discussed.

Chapter 2 : Designing Chatbots for Conversational Commerce—A Look into the Human-AI Conversation

Abstract

This study discusses how chatbot designs affect the human-chatbot interaction process and, thus, influence human-AI interaction outcomes. Chatbots are gaining momentum in a variety of business functions, such as IT service, human resource management, customer service, marketing, and sales. We are now entering a new era of conversational commerce with the great potential to save costs, improve efficiency, and enhance customer relationships in different business sectors. While organizations have the opportunity to leverage the benefits of chatbots, they also face challenges such as customers confusion and resistance from customers. Although chatbot proactivity design and social identity design are popular in the industry, limited existing research on chatbots has investigated chatbot proactivity design and its role as a boundary condition of chatbot social identity design. In order to optimize user experience with chatbots, there is thus a need from both the business and the academia aspects to better understand how chatbot social designs affect the human-chatbot conversational process and, in turn, affect user perception and task success from the perspective of uncertainty reduction. The field experiment results reveal that proactivity design will decrease communication inefficiency and, thus, increase customer satisfaction and the probability of task success. Furthermore, chatbot social identity design weakens the negative effect on customer satisfaction made by ineffective communication.

2.1. Introduction

Artificial Intelligence (AI) applications have been widely adopted in industries such as healthcare, finance, and marketing. While AI applications were focusing on supercomputing capabilities, they have evolved into people's daily life with distinctive features from previous technologies (e.g., computers and the Internet)—i.e., the 'natural' communication features and the anthropomorphic characteristics. The interface between AI and humans is no longer just standard commands, menus, or buttons that have steep learning curves for most people. Instead, AI can communicate with users using natural language. Moreover, anthropomorphic features such as personality and informal language bring social factors into human-AI communication that significantly improve user experience. One popular AI application is the chatbot technology which has been evolving for over 50 years. Recently, all leading technology companies—e.g., OpenAI, Google, Microsoft, Tencent, Amazon, Facebook, IBM, Apple, Samsung, etc.—have developed platforms to make chatbots simple and quick for business implementation. A few examples include ChatGPT from OpenAI (launched in 2022), Watsons from IBM (launched in 2011), Xiaoice from Microsoft (launched in 2014), and Amazon Alexa (launched in 2014).

Chatbots have gained significant traction in the business landscape, revolutionizing conversational commerce by providing a natural language user interface, particularly through voice or text interactions (Eeuwen, 2017). A notable example is Facebook's revelation at the F8 conference, which highlighted the presence of a staggering 300,000 active chatbots on their platform[1]. Among these advancements, ChatGPT has emerged as a noteworthy contributor to business success. Its implementation offers a wide array of advantages, including enhanced customer support, scalability, cost-effectiveness, swift response time, lead generation, sales support, market research facilitation, and feedback collection. In broader terms, chatbots can provide businesses with numerous benefits, such as expedited consumer service, reduced labor costs, multitasking capabilities, and consistent service quality (Intellexer, 2019). The chatbot market is expected to reach USD 102.29 billion, representing a 34.75 percent compound annual growth rate (CAGR) over the forecast period (2021-2026)[2]. However, despite the remarkable advancements in chatbots, even if it is ChatGPT, there are problems in the efficiency of human-chatbot communication (Agomuoh, 2023; Cem, 2023). It results from uncertainty level is high when people interact with a chatbot. For example, the behavioral uncertainty level is high when

[1] https://about.fb.com/news/2021/06/f8-refresh-developer-conference/
[2] https://www.mordorintelligence.com/industry-reports/chatbot-market

chatbots struggle to understand complex or ambiguous queries. For instance, if a user seeks the chatbot's assistance in finding a good restaurant nearby, the chatbot may fail to comprehend the user's specific preferences, resulting in inappropriate restaurant recommendations. Consequently, users experience frustration due to the inadequacy of the chatbot in addressing their needs accurately. In addition, users may also contribute to cognitive uncertainty by not providing straightforward and specific instructions or queries. When users input a non-specific prompt such as "help" without specifying the issue they require assistance with, the chatbot may deliver generic or ineffectual responses. This ambiguity further intensifies user frustration as their actual requirements may remain unaddressed due to the chatbot's limited understanding of their ambiguous input. Another aspect of cognitive uncertainty emerges when users encounter difficulty explaining complicated needs or requests to a chatbot such as customization. Users may struggle to express their preferences accurately or locate the appropriate phrases or choices within the chatbot interface. Consequently, communication inefficiency prevents the chatbot from providing the intended result, ineffective responses, leading to user dissatisfaction. Addressing and mitigating these uncertainties is crucial to improving the effectiveness and usability of chatbot systems, ensuring enhanced user experiences and outcomes.

Different chatbot design features have been widely applied in industries to tackle the human-chatbot communication problems such as empathy expression, personality, intimate tone, proactivity, social identity, etc. In this research, we mainly focus on proactivity design and social identity design to reduce uncertainty between human-chatbot communication. Chatbot proactivity involves both the initiation of conversation and providing multiple options for consumers to choose from in the chat (Peng et al., 2019). Proactivity features (e.g., a prompt, a welcome message) are expected to optimize user experience, increase conversions, and anticipate consumer needs[3]. Another example is to incorporate social features (e.g., giving a

[3] https://www.inbenta.com/en/blog/proactive-vs-reactive-chatbots-pros-and-cons/

certain human name to a chatbot) in chatbot design. According to a survey (Leah, 2018), eight of ten chatbots have human names, such as Rosie, Alfred, and Ruby, in the most popular chatbots. Practitioners believe that social identity design could make chatbots appear more approachable and help consumers feel more comfortable[4]. Though different design solutions are developed, there is mixed evidence about chatbot performance[5]. In addition, few studies have focused on the impact of these designs on the conversational process, and usually they skip right over the conversational process to study the business outcomes directly. Thus, managers have the pressing need to better understand the research questions—how human-chatbot communication inefficiency is influenced by the two chatbot social designs and ultimately, how the business outcomes are influenced. In the industry, conversion rate and user experience are the critical metrics of chatbot performance. In this research, we incorporate the industry metrics. We use the probability of task-oriented human-AI interaction success to represent conversion rate, and consumer satisfaction with the interaction represents user experience.

Chatbots commonly serve as the primary interface for customers engaging with businesses. However, ineffective communication between users and chatbots often leads to frustration and resistance, which hampers the realization of the significant value that chatbots can bring. Consequently, there is a need to thoroughly investigate and understand the issues about human-chatbot communication inefficiency. This research aims to delve into the human-chatbot conversational process, specifically focusing on the human-chatbot communication inefficiency. To evaluate the human-chatbot communication inefficiency, a novel metric is proposed as an essential measure. We conceptualize four types of human-chatbot communication inefficiency based on Li's study (Li et al., 2020) and define it as excessive cost in human-chatbot communication based on Vetter's definition (Vetter, 2000). We use the

[4] https://www.soocial.com/chatbot-names/
5 https://www.cmswire.com/consumer-experience/we-know-chatbots-are-falling-short-but-why/

communication inefficiency because inefficiency can better reflect the phenomenon and matches our measurement. According to existing literature (Li et al., 2020) and our human-chatbot conversation logs, we have summarised human-chatbot communication inefficiency into four types—(1) misunderstanding between consumers and chatbots, (2) conversation diverging into irrelevant topics, (3) asking the other party to repeat or rephrase questions, and (4) guessing the meaning of expressions. Furthermore, the study evaluates the impact of proactive and social identity design strategies on the communication inefficiency between humans and chatbots and, accordingly, how human-chatbot communication inefficiency affects customer satisfaction and task completion. In particular, we draw from the perspective of uncertainty reduction and examine the impacts of two design features—chatbot proactivity and social identity—on the inefficiency of human-chatbot conversation, which in turn influences user satisfaction and task success. We argue that chatbot proactivity design will reduce cognitive uncertainty by initiating the conversation with a set of potential topics. Thus, consumers will have a basic understanding of the talking scope and know how to ask questions to the chatbot. The social identity design will reduce behavioral uncertainty by setting social norms. Chatbot social identity design will indicate what behaviors are expected in the situation. In addition to the direct effect of social identity design, social identity design also has a moderation effect. Social identity will weaken the negative effect of communication inefficiency on consumer satisfaction with the interactions due to ingroup favoritism. People are usually more patient and friendly with their ingroup members (Taylor and Doria, 1981). Ingroup members are more tolerant of communication inefficiency and reduce the perceived frustration made by the communication inefficiency. As a result, ingroup human-chatbot communication has better satisfaction with the interaction. Consumer satisfaction is an affective outcome of the conversation, which is not only influenced by the communication inefficiency but also the consumer emotion bias. Therefore, social identity design has a moderation effect. It affects consumer satisfaction with human-AI interaction by two ways.

The first one is the direct effect from the perspective of behavioral uncertainty reduction. The second is the moderation effect caused by ingroup bias. Compared to consumer satisfaction with the interaction, task success is a representation of uncertainty reduction outcome, which is only influenced by communication inefficiency in this study. Therefore, we argue that social identity moderates the relation between communication inefficiency and consumer satisfaction.

The intended contribution of this research is to advance the study of chatbots within the Information System literature. Firstly, the research recognizes the need to address the problem of communication inefficiency in human-chatbot interactions in order to fully harness the potential value of chatbots. To this end, the research proposes a measure of communication inefficiency that can serve as a foundation for improving communication effectiveness in human-chatbot interactions. Secondly, this research seeks to bridge the gap in the existing literature on chatbots in the field of information systems. While most studies have extensively discussed the business influences of chatbot social designs (i.e., Schanke et al., 2021; Schuetzler et al., 2020), there has been a lack of research that delves into the specific nuances of the conversational process involved in human-chatbot interactions. Lastly, the research expands upon the utilization of uncertainty reduction research in the context of human-chatbot communication inefficiency studies. While previous research has used this theory to explore the distinction between human-human communication and human-chatbot communication, as well as its impact on friendship formation between chatbot and human (Lurings, 2019; Bahtiri and Peeters, 2019), there is a dearth of research that applies this theory to guide the reduction of human-chatbot communication inefficiency, despite its prevalence in contemporary communication practices. By incorporating uncertainty reduction research into the examination of communication inefficiency, this research aims to address a pressing issue in human-chatbot communication.

2.2. Literature Review

We conducted a review of the literature on chatbot in the fields of information system, human-computer interaction, and marketing. We identify two relevant streams of literature—one is business research on the anthropomorphic (humanness) and social design of chatbots and the other computer science research on human-chatbot conversation analysis.

Chatbot design literature mainly focuses on how anthropomorphic (humanness) and social features of chatbots—e.g., modality (voice vs. text) (Cho et al. 2019), chatbot disclosure (e.g., Luo et al., 2019; Shi et al., 2020), conversation style (i.e., Schuetzler et al., 2020; Kull et al., 2021), and personality (i.e., Jin et al., 2019; Tärning and Silvervarg, 2019)—influence user perceptions, intentions, and related business outcomes (Table 2.1). User perceptions include attitudes towards chatbots (i.e., Ashktorab et al., 2019; Cho, 2019; Tärning and Silvervarg, 2019), social presence (i.e., Gnewuch et al., 2018; Roy and Naidoo, 2021), perceived humanness (i.e., Candello et al., 2017; Araujo, 2018; Schuetzler et al., 2020), etc. Most research revealed positive relationships between anthropomorphic and social features of chatbots and user perceptions. For example, Beattie et al. (2020) found that messages using emoji is increased users' perceptions of the message source as more socially attractive, credible, and with more CMC competence. Business outcomes include purchase (i.e., Luo et al., 2019; Schanke et al., 2021), donation persuasion (i.e., Shi et al., 2020), and brand engagement (i.e., Kull et al., 2021). Again, most research revealed positive relationships between anthropomorphic and social features and the business outcomes. For instance, Schanke et al. (2021) found that the chatbot humor design and communication delays increased conversion. While all the features were postulated to produce a smooth conversation between users and chatbots, metrics of the conversations were not included in the research models—the mediators in the studies were user perceptions such as perceived humanness (e.g., Cho et al. 2019) and social presence (e.g., Lee et al., 2006; Schuetzler et al. 2020). The link between chatbot design and the conversation quality is missing in this stream of literature. Therefore, we propose communication inefficiency to represent conversation quality in this research by measuring the number of the occurrence of conversation non-smoothness—non-progress (Li et al., 2020) and obstacles (Myers et al., 2018)—divided by the number of conversations turns.

Research on human-chatbot conversation analysis takes advantage of natural language processing (NLP) techniques (e.g., Akhtar et al., 2019; Candello et al., 2019; Myers et al., 2018), existing AI services (e.g., IBM Watson service (Li et al., 2020)), or manual coding (e.g., Jain et al., 2018; Logacheva et al., 2018) to explore human-chatbot conversations (Table 2.2). Most of the research consists of exploratory and descriptive studies such as identifying conversation topics (e.g., Akhtar et al., 2019), user intents (e.g., Candello et al., 2019), and the differences between human-human and human-chatbot conversations (e.g., Hill et al., 2015; Logacheva et al., 2018). In particular, a group of studies examined human-chatbot communication (in)efficiency, such as identifying non-progress types (e.g., Li et al., 2020; Myers et al., 2018;), user coping strategies (Candello et al., 2019; Li et al., 2020; Myers et al., 2018), and chatbot repair strategies (e.g., Ashktorab et al., 2019). However, most of the literature does not investigate the effects of chatbot design on human-chatbot conversations. Also, the existing computer science studies are not focusing on business contexts and thus do not examine the links between conversation analysis and business outcomes such as consumer decision making and satisfaction.

In summary, while a variety of chatbot design features have been examined to influence different business-related conversation outcomes, existing business literature has limited understanding of the conversation as the mechanism underlying the relationship between chatbot designs and the outcomes. On the other hand, though computer science research has analyzed the human-chatbot conversations, a better understanding is needed of how the efficiency of such communication is linked to both chatbot design and outcomes that are relevant to the business context. In the current research, we integrate these two streams of research to investigate the effects of chatbot design on the conversation that in turn affects user satisfaction and task outcome. In particular, drawing from the literature on conversation analysis, we identify two relevant notions—non-progress (Li et al., 2020) and obstacles (Myers et al., 2018)—as a conceptual foundation of human-chatbot communication inefficiency, which we theorize to influence user satisfaction and task success. Moreover, following the distinction between cognitive uncertainty and behavioral uncertainty in uncertainty reduction theory in communication research (Berger and Calabrese, 1975)), we aim to investigate two chatbot design features—proactivity and social identity—that affect communication inefficiency. We choose these two features because they respectively address the two types of uncertainty—proactivity for cognitive uncertainty reduction and social identity for behavioral uncertainty reduction. We discuss the theoretical foundation and develop our hypotheses in the next section.

Table 2.1. Business Research on Chatbot Design

Source	Design-related IVs	Mediator	Moderator	DVs	Findings
Anthropomorphic (humanness) Features					
Araujo (2018)	Human-like design (informal language and name) vs. machine-like design			Mindful and mindless anthropomorphism	In the consumer service context, subjects were required to fulfil a task of changing order addresses with a chatbot. Results suggest that human-like agents increase users' perceptions of mindful and mindless anthropomorphism.
Cho et al. (2019)	Modality (voice vs. text)	Perceived human-likeness	Task type (utilitarian vs. hedonic)	Positive attitude towards chatbots	Participants were required to communicate with chatbots in a lab experiment to complete a task (utilitarian or hedonic). 1. The paper revealed a positive effect of laptop usage and on the perceived human likeness of chatbots. 2. The positive effect of voice interaction on attitudes toward chatbot is mediated by perceptions of human-like characteristics in utilitarian task conditions.
Candello et al. (2017)	Typeface type (human-like vs. machine-like)			Users' perception of humanness	In the lab experiment, participants communicated with a financial advisor chatbot. It is revealed that while machine-like typeface was perceived more as a machine, the human-like typeface was *not* perceived more as human.
Gnewuch et al. (2018)	Dynamic delay vs. non-dynamic delay according to the complexity of user questions			Perceived chatbot humanness, social presence, and satisfaction	The results indicate that dynamic response delays can increase users' perception of chatbot humanness, social presence, and satisfaction in the context of consumer service.
Luo et al. (2019)	Chatbot identity disclosure		Chatbot identity disclosure timing	Purchase	In a sales call setting, researchers conducted a field experiment and found that: 1. Chatbots identity disclosure will reduce the purchase likelihood.

Source	Design-related IVs	Mediator	Moderator	DVs	Findings
					2. The above negative effect will be weakened by late chatbot identity disclosure.
Adam et al. (2020)	Anthropomorphic design cues (chatbot identity, small talk, and empathy design)	Social presence		User compliance	An online experiment was conducted in the context of banking. The results indicate that anthropomorphic cues (chatbot identity, small talk, and empathy design) increase user compliance mediated by social presence.
Social Features					
Smestad and Volden (2019)	Chatbot personality (agreeable vs. conscientious)			User experience measured by AttrakDiff evaluation	This study found that agreeable personality had a significant positive effect on the user experience of chatbots.
Roy and Naidoo (2021)	Conversation style (warm vs. competent)	Social perception of the brand	Time orientation (present vs. future)	Attitudes towards brand and purchase intention	Participants joined a simulation of shopping on a hotel website. Consumers' time orientation (present orientation) strengthens the positive effect of warm style conversation on attitudes towards chatbot and purchase intention.
Schuetzler et al. (2020)	Conversational skill (tailored responses and response variety)	Social presence		Perceived humanness and partner engagement	The participants were required to communicate with chatbots in an online experiment. The research indicates that individuals perceive a chatbot with high communication skills to be more socially present, which leads to a positive impact on perceived humanness and partner engagement.
Beattie et al. (2020)	Chatbot message type (with emoji vs. without emoji)			Social attraction, CMC competence, and the credibility of message source	In a restaurant recommendation context, researchers found that messages using emojis increased users' perceptions of the message source as more socially attractive, credible, and with more CMC competence.
Kull et al. (2021)	Initial message tone (competent vs. warm)	Brand-self distance	Brand affiliation (consumer vs. non-consumer)	Brand engagement	The study was conducted in the context of travel guide and banking, respectively. Results revealed:

Source	Design-related IVs	Mediator	Moderator	DVs	Findings
					1. Brand engagement increases if chatbots can initiate conversation using a warm welcoming message. 2. Brand–self distance mediates the above effect, such that a warm (vs. competent) initial message makes consumers feel closer to the brand. Further, brand affiliation strengthens the relationship between message warmness and brand engagement.
Anthropomorphic and Social Features					
Schanke et al. (2021)	Social presence (intimate tone), communication delays (typing time), and humor (joking) designs		Cash offer	Transaction conversion	In the context of online retailing, the results indicate: 1. The anthropomorphic designs increase conversion, single anthropomorphic design increase conversion by 6.7%, all three anthropomorphic designs increase conversion by 10.8%. 2. This result suggests that customers become much more price-sensitive in the presence of enough anthropomorphism.
Shi et al. (2020)	Chatbot identity disclosure		Persuasion strategies (personal vs. non-personal)	Donation persuasion	The research revealed that: 1. If users perceive the partner as a human, they are more likely to donate. 2. If the users perceive the partner as a human and receive personal inquiries persuasion strategy, the persuasion effect is stronger.

Table 2.2. Computer Science Research on Human-chatbot Conversation Analysis

Source	Findings
Exploratory and Descriptive Research	
Akhtar et al. (2019)	The human-chatbot conversation log was analyzed with a Cross-Industry Standard Process for Data Mining model(CRISP-DM reference model). The research found that users were generally interested in and satisfied with their conversations with chatbots. If users model not receive instant and expected responses to their inquiry, they would quit the chats.
Park et al. (2018)	Researchers analyzed the conversation log between students and two chatbots (Rose and Mitsuku). They found that students used simple language and made text errors. Other findings were that Rose seemed to switch topics more often while Mitsuku seemed to be more combative in conversation.
Logacheva et al. (2018)	Human-human and human-chatbot conversations were compared to find the differences between the two—e.g., human-human conversations had more diverse topics and longer durations.
Hill et al. (2015)	The research compared human-human conversations with human-chatbot conversations by an NLP method. Researchers analyzed words per message, message per conversation, profanity, emotion, and so on. The results revealed that humans used fewer words per message, more messages per conversation, and more profanities towards the chatbot as compared to human-human conversations.
Jain et al. (2018)	Researchers analyzed human-chatbot conversations by manual coding. They found users liked chatbots that could talk in natural language in a 'human-like' manner or ones that engaged users.
Human-chatbot Communication Inefficiency	
Li et al. (2020)	In a digital-banking context, 17136 conversations were both manually coded and analyzed by AI—IBM Watson conversation understanding service. The research identified 12 types of "non-progress" (NP) of the conversations and 10 user coping strategies such as switching topics, reformulation of questions, and quitting. The research found that users had the intention to quit if they had encountered 3 consecutive incidences of NP. Moreover, Reformulation was used as the last strategy to cope with NP.
Candello et al. (2019)	Conversation logs and videos between users and chatbots were analyzed with a semi-supervised topic cluster methodology to identify 4 topics from the conversations. The results suggest that the presence of an audience in the conversation increased users' reaction to chatbot failures.
Myers et al. (2018)	The study used Dialogflow to analyze voice conversation records between humans and a chatbot. The results identified 4 types of conversation obstacles, including unfamiliar intent, NLP error, system error, and failed feedback. The study then revealed user coping tactics such as simplification, quitting, and recall.

Source	Findings
Ashktorab et al. (2019)	By conducting paired comparison experiments on M-Turk, the research examined users' preferences for chatbot repair strategies (top response, repeat, confirmation, options, defer, keyword highlight explanation, keyword confirmation explanation, and out-of-vocabulary explanation). 1. It revealed that the option repair strategy is the most favorite strategy while the repeat strategy led to the most unsuccessful tasks. 2. When the repair was effective, defer strategy rated last, but when the repair failed, defer ranked second.
Corti and Gillespie (2016)	In a student sample experiment, results suggest that: 1. When the chatbot was implemented through a text interface, and users knew their interlocutor was a chatbot, users were less likely to initiate repairs. 2. When the chatbot was implemented through a human-body interface to simulate face-to-face human-human communication, users made the greatest inter-subjective effort to build common ground for the conversation.

2.3. Theoretical Foundation

We propose uncertainty reduction theory as a theoretical foundation. The theory postulates that people dislike uncertainty in their interpersonal interactions (Berger and Calabrese, 1975). In the situation of an initial interaction, uncertainty is high because individuals are unaware of the other party's opinions and attitudes (Berger, 1982). There are two types of uncertainty in communication—cognitive uncertainty and behavioral uncertainty. Cognitive uncertainty refers to the uncertainty about the other party's beliefs and thoughts, while behavioral uncertainty refers to the failure to explain or predict the other party's behavior or failing to know what behaviors are expected in the situation (i.e., social norms) (Redmond, 2015). Both types of uncertainty are of great importance to efficient communication. A high level of uncertainty will lead to misunderstandings in communication (Paek and Horvitz, 1999) and a negative state of discomfort and anxiety (Deng et al., 2021). People are thus motivated to eliminate uncertainty in communication with uncertainty reduction strategies (Berger, 1982). For example, Berger and Calabrese (1975) proposed three types of such strategies—passive strategy (e.g., observing the targeted person), interactive strategy (e.g., directly asking questions or self-disclosure to the targeted person), and active strategy (e.g., searching for information about the targeted person). Uncertainty reduction theory has been applied in intercultural communication research (i.e., Nelson, 1992), job hiring process research (i.e., Ragan, 1983), and computer-mediated communication research (i.e., Pratt et al., 1999; Antheunis et al., 2011). The concept of proactivity has been studied extensively in the fields of organizational behavior (i.e., Grant and Ashford, 2008) and psychology (i.e., Wanberg and Kammeyer-Mueller, 2000). Proactive individuals are known to exhibit traits such as problem-solving orientation, taking initiative, and seeking opportunities for personal growth and development. Such proactive behaviors have been associated with reduced cognitive uncertainty as individuals actively seek and acquire information to increase their understanding of a given situation or task. Similarly, the concept of social identity has been extensively researched in the field of social psychology (Muldoon et al., 2019; Reicher et al., 2010). Social identity refers to the extent to which individuals perceive themselves as being part of a larger social group or category, and this identity can influence behavior and decision-making. When individuals strongly identify with a social group or have a well-defined social identity, it can provide a sense of belonging and reduce behavioral uncertainty by offering clear norms and guidelines for behavior within the group.This research further extends the theory to human-chatbot communication. When people are engaged in an initial interaction with a chatbot, uncertainty can be even higher than communicating with another human because people lack an understanding of how chatbots work and are uncertain about the chatbots' capacity and communication style (Liu, 2021). Consequently, the communication between human users and chatbots can be inefficient. Generally, communication inefficiency refers to excessive cost for sending, receiving, or transferring messages (Vetter, 2000). In human-chatbot communication, it is manifested in the obstacles (Myer et al., 2018) or non-progress of human-chatbot conversation (Li et al., 2020). For example, human users may talk about out-of-scope topics, produce typos and unfinished messages, and split one message into multiple utterances in their conversation with chatbots, which causes chatbots to misunderstand users' messages. On the other hand, chatbots may have a low accuracy of user intent capture, which holds up the conversation until extra explanations are provided by human users.

Following the notion of uncertainty reduction, the current research examines two features of chatbot design—proactivity and social identity—as the uncertainty reduction strategies that reduce human-chatbot communication inefficiency, which in turn influences user satisfaction and task success in the context of consumer product customization (Figure 2.1).

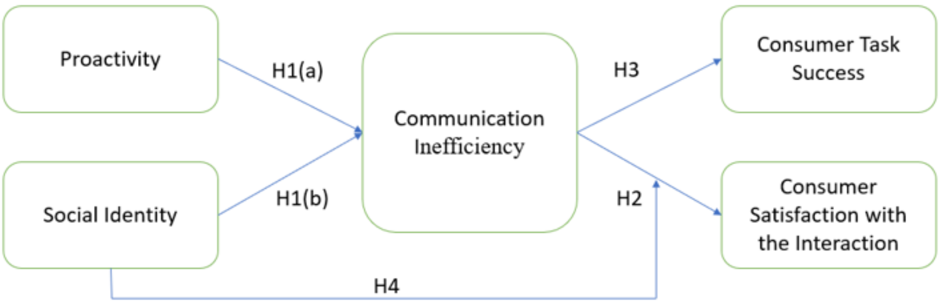

Figure 2.1. Research Model

2.4. Hypothesis Development

We hypothesize that proactivity design will reduce communication inefficiency by reducing user cognitive uncertainty about conversation initiation and the scope of the chat. The proactivity of chatbots involves the initiation of a conversation and providing multiple options for consumers to choose from (Peng et al., 2019). In the context of conversational commerce, chatbot proactivity is a form of *prompt* strategy—i.e., chatbots initiate the conversation with a welcome message and provide a list of options for consumers to choose, such as "Do you want to (1) … (2) … (3) …?". Consumers tend to feel uncertain about how to start a conversation with chatbots and describe the conversation topic(s) at the beginning of chatting with bots (Li et al., 2020). Chatbot proactivity design can reduce such uncertainty by starting the conversation with a set of topics of potentially relevant to consumer intentions (David, 2016). Thus, consumers will talk less about out-of-scope topics, which leads to fewer misunderstandings and reduced communication inefficiency. Research has found that chatbot proactivity design can provide additional information, improve bot productivity, guide consumers, and better frame a conversation to reduce communication inefficiency (Chaves and Gerosa, 2020). Therefore,

H1(a): Chatbot proactivity design will have a negative effect on communication inefficiency.

Chatbot social identity design will reduce communication inefficiency by reducing users' behavioral uncertainty in their conversations with bots. Social identity refers to an individual's self-concept drawn from perceived membership of a social group (Turner and Oakes, 1986) and indicates the expected proper behavior in the group (Katzenstein, 1996). In human-human communication, social norms generally guide the conversation process. However, in human-chatbot communication, they lack such norms. By assigning a social identity to a chatbot—e.g., giving the chatbot a human name that consumers are familiar with and changing the chatbot's conversation style to be more intimate—social identity design sets up the norms of a human-chatbot conversation, which in turn reduce consumers' behavioral uncertainty during the conversation and thus improve communication efficiency. Existing literature also found that the social presence of chatbots led to more social interpretations of the conversation from consumers and increased the likelihood of successful transactions (Schanke et al., 2021). Thus, the social identity design can help consumers have a better understanding of the behavioral norms and reduce inefficient communication in a human-chatbot conversation.

H1(b): Chatbot's social identity design has a negative effect on communication inefficiency.

Communication inefficiency will negatively influence consumer satisfaction. Satisfaction generally refers to "the affective response to the fulfillment of expectation-type standards" (Hecht, 1978, p.350). Inefficient communication, such as misunderstandings, hinders communication progress and leads to frustration (Robles, 2017). Therefore, ineffective communication has a negative impact on interactants' satisfaction. This negative relationship has been documented in many contexts such as education (the communication between instructors and students (Sidelinger et al., 2015)), healthcare (the communication between patients and doctors (Burgener, 2020)), and human resource management (the communication between employees and employers (Jacobs et al., 2016)). In the consumer context, communication inefficiency, such as misunderstandings and irrelevant topics, will lead to consumers' frustration because they cannot achieve their goals—e.g., inquiry about product information or receiving after-sale services—despite their attempts in the conversation to have chatbots understand their intents (Weidemann and Rußwinkel, 2021). This poor experience will decrease consumer satisfaction. Therefore,

H2: Communication inefficiency has a negative effect on consumer satisfaction.

Communication inefficiency will negatively affect the probability of task success. Ineffective human-chatbot communication disrupts uncertainty reduction and may even produce misinformation, which cause the task to fail. In the IS research, ineffective communication has been found to have a negative effect on the success of information system development (Edstrom, 1977). In our research context, consumers may feel more confused and uncertain about how to communicate with chatbots to complete a task in the human-chatbot conversation. Such a high level of uncertainty will lead consumers to abandon the task or fail to complete the task due to misinformation and/or a lack of information. Existing studies find that consumers are likely to quit the conversations when human-chatbot communication is not effective (Li et al., 2020). Therefore,

H3: Communication inefficiency has a negative effect on the probability of consumer task success.

Besides the direct effect on communication inefficiency by setting shared norms, chatbot social identity will also moderate the relationship between communication inefficiency and consumer satisfaction. We believe that chatbot social identity will moderate the negative relation between communication inefficiency and consumer satisfaction with the interaction due to ingroup favoritism. Social identity is a popular chatbot design used to shorten social distance between chatbots and human users. Social identity is "part of an individual's self-concept which derives from his knowledge of his membership of a social group (or groups) together with the value and emotional significance attached to the membership" (Tajfel, 1974, p.49). Social identity theory explains how social groups create a positively distinctive identity (Tajfel and Turner, 1979). This distinctive positive identity leads to ingroup favoritism, where ingroup members are treated more favorably (Taylor and Doria, 1981). When a consumer observes a familiar and close social identity with a chatbot, he/she will have the ingroup favorableness toward the chatbot and be more patient, accommodating, or forgiving when encountering communication inefficiency in the human-chatbot conversation. This will mitigate the negative effect of communication inefficiency on consumer satisfaction. Research has found that visual similarity between users and chatbots leads to users' perceived group identity and, thus, increases users' trust towards chatbots (Xu and Lombard, 2017). Therefore,

H4: Chatbot social identity weakens the negative effect of communication inefficiency on consumer satisfaction.

2.5. Method

2.5.1. Research Setting

We conducted a field experiment to test our hypotheses. The research setting was an online product-customization task in which participants were required to customize the package of a branded beer with their photos or pictures (see Figure 2.2). Product customization was chosen as the focal task because (1) it is a popular marketing strategy to generate more revenue and build consumer loyalty (Lu, 2017); and (2) consumers tend to need a consumer service agent's assistance with product customization because compared to purchasing standard products, customizing products requires more steps to complete and consumers maybe not be familiar with the product design process. Therefore, the beer brand implemented a chatbot to facilitate the customization task.

The brand we collaborated with was a popular beer brand in Southeast China. The focal brand has a social media community on the WeChat platform (one of the most popular Chinese social media platforms) with around 1.5 million subscribers. The company provided a customized beer product that is only for sale on WeChat. Thus, the brand's WeChat subscribers were the target subjects of our study, and we utilized this real business setting to test our research hypotheses.

Figure 2.2. Self-customized Beer Package

2.5.2. Manipulation

The experiment followed a 2 (proactivity: Yes vs. No) * 2 (social identity: Yes vs. No) between-subject design (Figure 2.3). The chatbot proactivity condition was manipulated by chatbot waiting for the subjects to initiate conversation with inputs into the chat window (the No condition) vs. chatbot starting off the conversation with greetings and two options for the subject to choose from—(1) to know more about the customized products; and (2) to tell the chatbot how the subject would like to customize the products (the Yes condition). The chatbot social identity was manipulated by assigning different names to the chatbot and varying the degree of intimacy in the greeting messages and follow-up utterances. For the control group without social identification, the chatbot was given a common name "Xiaoming", while in the treatment group, the chatbot was named "Brother Yang"—the name of a virtual figure representing the brand in its WeChat community to interact with customers, such as regularly posting promotional contents and following up with customers' comments. This virtual figure is backed up by a human team and is a well-known social identity in the WeChat community—customers often directly address the virtual figure as "Brother Yang" in their comments. All members of the community were familiar with this virtual identity, and this setting served as the foundation of our manipulation of social identity. That is, we leverage the established social bond between the virtual figure—"Brother Yang"—and the community members to assign the social identity to the chatbot in the treatment group.

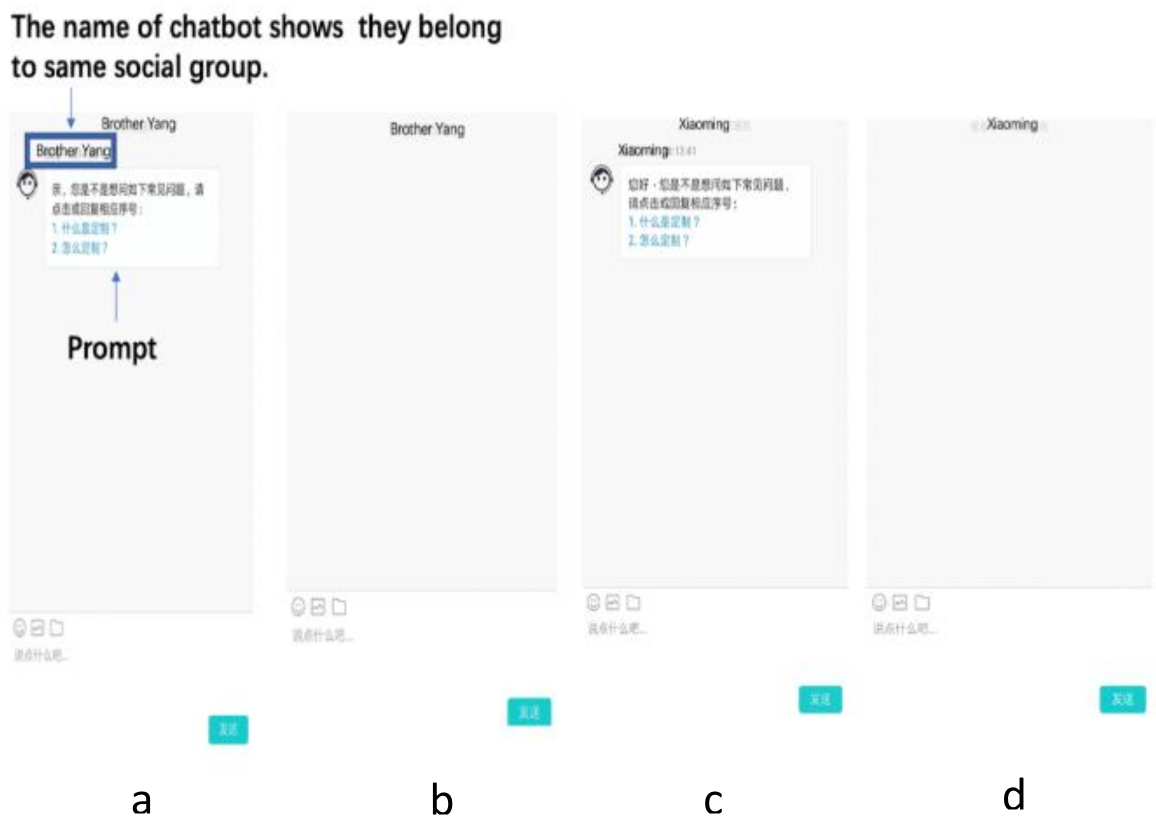

Figure 2.3. a-d Screenshots of the Manipulations in the Chatbot Chat Windows

Note: Figure 2.3 a is the condition with the proactivity and social identity manipulation; 2.3 b is the condition with the non-proactivity and social identity manipulation; 2.3 c is the condition with the proactivity and non-social identity manipulation. 2.3 d is the condition with the non-proactivity and non-social identity manipulation.

2.5.2.1 Manipulation Check

To ensure that customers were appropriately experiencing the two treatments, we conducted a manipulation check after the online field experiment. 30 participants were recruited for cash incentive. Participants were required to read the four types 2 (proactivity: Yes vs. No) * 2 (social identity: Yes vs. No) of consumer-chatbot interaction interfaces (randomly ordered) and answered the manipulation check questions. For social identity treatment, we asked participants to assess how strongly they agreed with the statements on a 7-Likert scale from strongly disagree (1) to strongly agree (7). The statement is that customer service demonstrated the social identity. For proactivity treatment, we asked the question: in the interaction, the customer service is proactive with the answer yes (1) or no (0). We analysed the data with a T-test. The result revealed significant differences in the proactivity condition and non-proactivity condition and the social identity and non-social identity condition ($p<0.01$) (See Table 2.3).

Table 2.3. Manipulation Check

Group	Mean	N	p-value	t-value
Proactivity vs. Non-proactivity with social identity	0.933 vs. 0.367	30	P<0.01	6.158
Proactivity vs. Non-proactivity with non-social identity	0.967 vs. 0.367	30	P<0.01	6.596
Social identity vs. Non-social identity with proactivity	4.767 vs. 3.600	30	P<0.01	4.364
Social identity vs. Non-social identity with non-proactivity	4.600 vs. 3.337	30	P<0.01	4.570

2.5.3. Procedure

The field experiment was launched in April 2019 on the homepage of the beer company's WeChat community[6]. We utilized the company's mobile system to randomly assign subjects into different experiment groups and record their micro-behaviour in the experiment such as landing time, leaving time, and chat logs. All subjects entered the system through the WeChat page of the brand community and finished the tasks and steps of the experiment on their mobile phone. All participants were members of the WeChat community, and they were incentivized by a lottery to join a promotion campaign for product customization. After joining the campaign, participants were told that the company was promoting the customized beer product and they would need to design a customized package by communicating with a chatbot agent. The experiment consisted of four steps (Figure 2.4). The first step was a short survey of participant demographics. Then a briefing about the customization task was presented. In the task briefing, subjects were told to finish the task with the assistance of a newly introduced chatbot agent. After the briefing, participants were randomly assigned to different experimental groups to finish the customization task (Figure 2.3 a-d). After that, a survey was conducted with the measurement instrument for the variables in our research model.

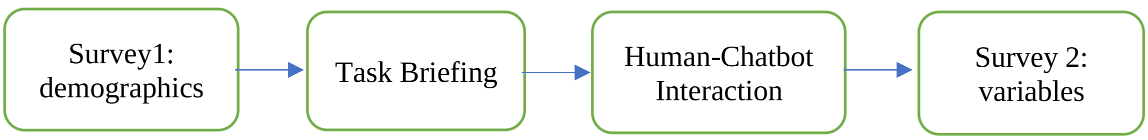

Figure 2.4. The Experiment Procedure

2.5.4. Measurement Instrument

In Survey 1, we collected the subject's age, gender, income, and prior AI experience. In particular, age is measured in four age groups. Gender is measured by binary data (Female vs. Male). Prior AI experience is measured by asking subjects the question—"Have you ever

[6] WeChat is one of the most popular social media platforms in China.

interacted with AI applications? (Yes/No)". The answers are coded with 1 for Yes and 0 for No. In Survey 2, we measured subjects' satisfaction with their conversations with the chatbot. Task success is whether the participant sends the customized picture to the chatbot (Yes vs. No: 1 vs. 0). If the participant successfully sends the picture, the chatbot will respond with a clear reply indicating the successful completion of the customized task. The scale of interaction satisfaction is drawn from the satisfaction measurement (Spreng et al.,1996) and consists of 3 items with a 5-point Likert scale from very disagree (1) to very agree (5)— 1) I am very satisfied with the customization interaction process in the previous step; 2) I am very delighted with the customization interaction process in the previous step.; and 3) I am very pleased with the interaction process in the previous step.

As the introduction discussed, chatbots are often hindered by users (Luo et al., 2019). For example, users don’t know how to communicate with chatbots and write unfinished message and try to finish the unfinished message (Li et al., 2020). In addition, the inability of chatbots to recognize context frequently results in irrelevant or useless replies (Giannelis, 2023). These lead to the communication inefficiency, which is a significant barrier between users and chatbots. Therefore, this research delves into the conversational process and discusses the communication inefficiency between humans and chatbots. Because everyone has different chatting habits, some people may like to ask a lot of questions. Therefore, we do not directly use conversation turns as a measure of inefficiency. Instead, we use the count of inefficiency divided by the number of conversation turns. We measure communication inefficiency as the ratio of conversation excessive cost to the number of conversation turns. Excessive cost refers to the occurrence of non-progress (Li et al., 2020) and/or obstacles (Myers et al., 2018) in a human-chatbot conversation. Li et al. (2020) analyzed human-AI communication processes and identified two non-progress patterns (excessive cost)—mis-recognition and non-recognition. Based on these patterns, we identified and coded four types of excessive cost—misunderstanding between subjects and the chatbot (mis-recognition), diverging on irrelevant topics (mis-recognition), asking the other party to repeat or rephrase questions (non-recognition), and guessing the meaning of expressions (non-recognition) (see Table 2.4). When chatbots encounter an unrecognizable utterance, they prompt the user to rephrase their question “您能换个方式问吗?(Could you please rephrase your question?) ” (Rephrase question situation). When the chatbot encounters a statement that it cannot recognize but has a similar topic in the database, it guesses "您是在问 XXX 吗?(Are you asking for XXX?) " (Guess meaning situation). The assessment of communication inefficiency is conducted through a approach adhering to the following outlined procedures: First, an extensive review of the pertinent literature and human-chatbot communication logs is undertaken. This comprehensive literature review enables the identification and categorization of four distinct types of communication inefficiency. Second, a group of independent laborers is assigned the task of labelling these inefficiencies. Each labeller performs this labelling task independently. Third, three researchers independently evaluate the consistency of the labelling provided by the labellers. Any minor discrepancies are discussed among the researchers and resolved, leading to the final measurement of communication inefficiency.

Table 2.4. Types of Communication Inefficiency

Communication Inefficiency Type	**Example**
Misunderstanding (mis-recognition)	Consumer A: I like pictures with flowers, but I don't have any now. Chatbot B: You can send it to me when you have time.
Diverging to irrelevant topics (mis-recognition)	Consumer A: This customization reminds me of the online games I've been playing. Chatbot B: Ha, I also like playing games. What games do you like?

Communication Inefficiency Type	Example
Asking the other party to repeat or rephrase questions (non-recognition)	Consumer A: Let me think about which picture to use. Chatbot B: ok. Consumer A: em… maybe. Chatbot B: Could you rephrase your question?
Guessing the meaning of expressions (non-recognition)	Consumer: I want to customize a. Consumer: Good! Consumer A: Slogan. Chatbot B: Are you asking for what the slogan is?

2.6. Data Analysis

Due to the nature of beer products, 89% of participants are male. A total of 1020 valid responses were obtained. The average age is around 32 years old. The average monthly income is 5750 RMB (about 846 USD). The descriptive statistics and the validation of the measurement scales are shown in Table 2.5.

Table 2.5. Descriptive Statistics

variable	N	mean	sd	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Satisfaction	1020	3.770	0.744	1								
Task Success	1020	0.261	0.439	0.155	1							
Communication Inefficiency	1020	0.165	0.239	-0.057	-0.266	1						
Age	1020	2.646	0.831	-0.029	-0.18	0.086	1					
Gender	1020	0.104	0.305	-0.073	0.047	-0.011	-0.11	1				
Income	1020	2.401	0.756	0.061	0.028	0.027	0.15	-0.21	1			
Prior Experience	1020	0.637	0.481	-0.134	-0.137	-0.014	0.125	0.01	-0.102	1		
Proactivity	1020	0.500	0.500	0.022	0.143	-0.332	-0.015	0.019	0.022	0.045	1	
Social Identity	1020	0.506	0.500	0.042	0.015	0.011	-0.02	-0.075	0.029	-0.016	-0.02	1

2.6.1. Results

The research model was examined by structural equation modelling (SEM) using Mplus. Mplus can handle a combination of different types of variables (such as count data, continuous data, and categorical data) (Muthén and Muthén, 1998-2010). The model controlled age, gender, income, and prior AI experience with a loglikelihood of -2791.8. The model's AIC is 5649.6. The BIC is 5812.2. The SEM results indicate that chatbot proactivity design reduces communication inefficiency by 15.8% ($\beta=-0.158$, $p<0.01$), which supports H1(a). Chatbot social identity design does not significantly influence communication inefficiency (does not support H1(b)). The potential reason is that the brand's social media identity appears in a new context—customisation—and therefore does not help the customer with the communication. Communication inefficiency reduces consumer task completion intensely ($\beta=-3.290$, $p<0.01$), which supports H2. One unite increases in communication inefficiency reduces human-chatbot interaction task success by around 96% ($1-e^{-3.290}$). At the same time, one unit increases in communication inefficiency decrease 37.8% consumer communication satisfaction ($\beta=-0.378$ and $p<0.05$). Chatbot social identity design also weakens the above negative relationship ($\beta=0.354$ and $p<0.1$). The results support H3 and H4. Communication inefficiency mediates the positive relationship between chatbot proactivity and task success (indirect effect: $\beta=0.521$, $p<0.01$ and total effect: $\beta=0.815$, $p<0.01$). Meanwhile, communication inefficiency mediates the positive relationship between chatbot proactivity design and consumer satisfaction (indirect effect: $\beta=0.060$, $p<0.05$ and total effect: $\beta=0.081$ $p<0.1$).

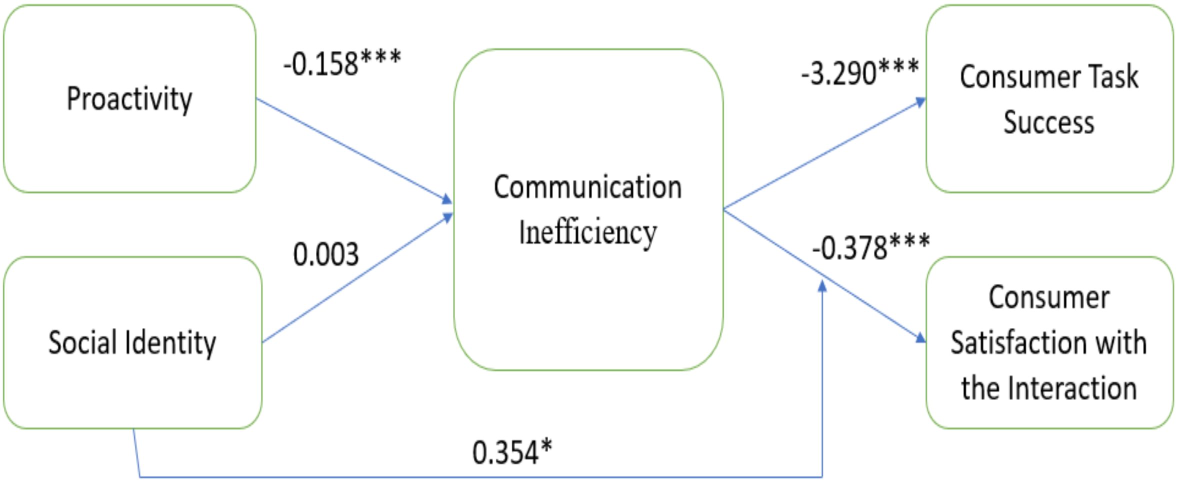

Figure 2.5. SEM Results

Table 2.6. SEM Results

IV \ DV	Communication Inefficiency (Negative Binomial Regression)	Consumer Task Success (Logistic Regression)	Consumer-chatbot communication Satisfaction
Covariates:			
Age	0.023**	-0.453***	-0.012
Gender	0.007	0.342	-0.157*
Income	0.007	0.193*	0.030
Prior AI Experience	-0.004	-0.616***	-0.199***
Independent Variables:			
Chatbot Proactivity	-0.158***	0.294*	0.021
Chatbot Social Identification	0.003	0.078	-0.004
Communication Inefficiency		-3.290***	-0.378**
Communication Inefficiency * SI			0.354*

*p<0.1; **p<0.05; ***p<0.01.

Table 2.7. Communication Inefficiency, Satisfaction and Success in Four Groups

Group	Communication Inefficiency Ratio 0-1		Satisfaction 1-5		Success 0 vs. 1	
	Mean	Std.	Mean	Std.	Mean	Std.
Non-proactivity and Non-social identity Group	0.242	0.756	3.680	0.767	0.178	0.383
Proactivity and Non-social identity Group	0.086	0.186	3.790	0.737	0.327	0.470
Non-proactivity and Social identity Group	0.246	0.262	3.818	0.749	0.217	0.413
Proactivity and Social identity Group	0.085	0.182	3.783	0.720	0.320	0.468

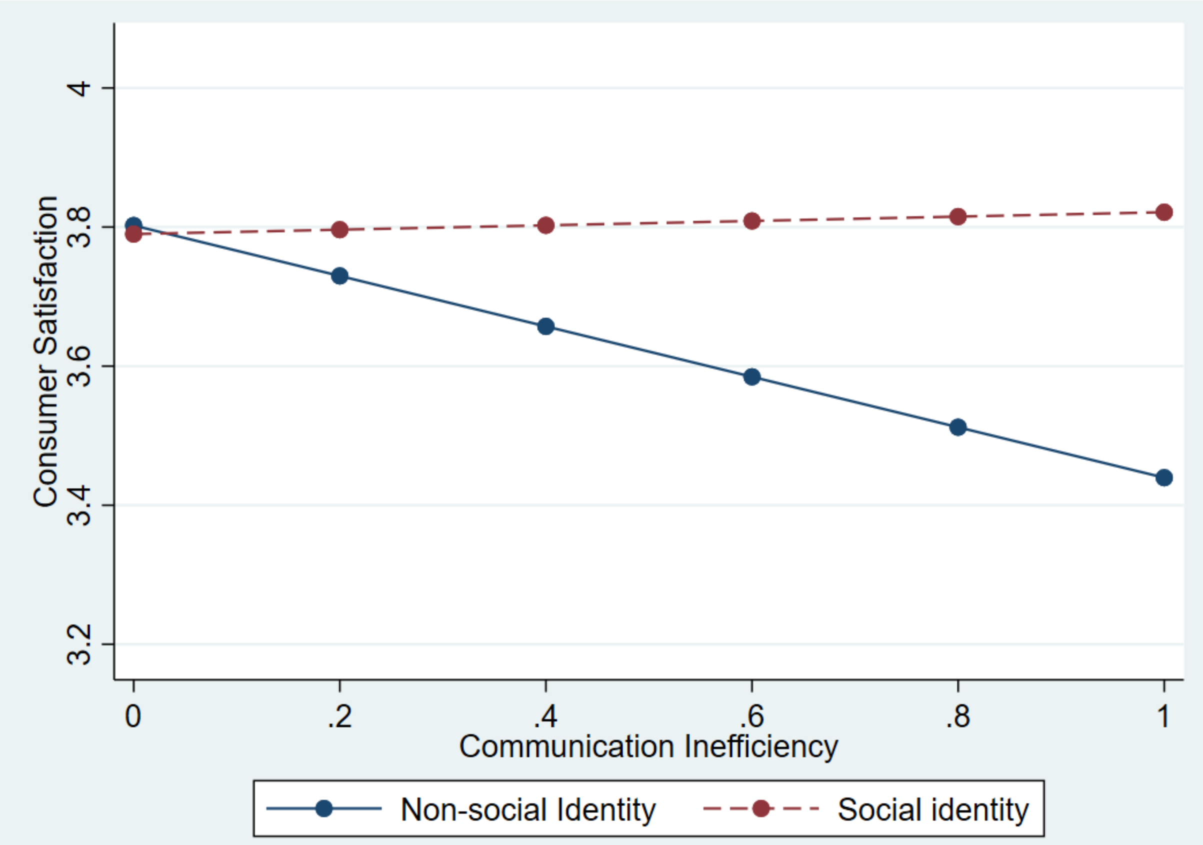

Figure 2.6. Moderation Effect

2.6.2. Robustness Check and Randomization Check

We further validated our research model by two robustness checks. First, we applied the stepwise regression method to check the robustness of the model. It indicates the results are robust to the findings (see Table 2.8). Second, we controlled the duration of the consumer-chatbot interaction to rule out the confounding effects made by interaction duration (in seconds). The results are consistent with the main findings (see Table 2.9), providing further support for our research model.

In addition to the robustness check, we also conducted randomization check with four demographic variables according to F-test statistics (see Table 2.10). Almost no significant differences were found among the four treatment groups (p>0.1). The significant differences in gender results from the gender imbalance (89% of participants are male).

Table 2.8. Stepwise Regression Results

Variables	(1) Inefficiency	(2) Success	(3) Ln(Satisfaction)	(4) Ln(Satisfaction)
Control Variables:				
Age	0.142***	-0.453***	-0.001	-0.001
	(2.577)	(-4.661)	(-0.093)	(-0.110)
Gender	0.0533	0.342	-0.044*	-0.044*
	(0.373)	(1.359)	(-1.746)	(-1.734)
Income	0.047	0.193*	0.014	0.014
	(0.773)	(1.859)	-1.431	-1.519
Prior AI Exp	-0.022	-0.616***	-0.058***	-0.058***
	(-0.246)	(-3.888)	(-4.206)	(-4.183)
Independent Variables:				
Proactivity	-1.048***	0.294*	0.004	0.004
	(-9.877)	(1.830)	(0.291)	(0.295)
Social identity	0.017	0.078	0.015	-0.006
	(0.193)	(0.505)	(1.070)	(-0.356)
Inefficiency		-3.290***	-0.066**	-0.129**
		(-8.088)	(-1.975)	(-2.542)
Inefficiency*Social identity				0.124*
				(1.876)
Constant	0.133***	0.175	1.318***	1.326***
	(4.879)	(0.469)	(37.019)	(36.720)
Observations	1,020	1,020	1,020	1,020
(pseudo) Adjusted R-squared	0.050	0.118	0.024	0.028

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.9. Robustness Check Results

Variables	(1) Inefficiency	(2) Success	(3) Ln(Satisfaction)	(4) Ln(Satisfaction)
Control Variables:				
Age	0.126**	-0.536***	-0.004	-0.004
	(0.055)	(0.100)	(0.009)	(0.009)
Gender	0.018	0.289	-0.047*	-0.047*
	(0.147)	(0.262)	(0.025)	(0.025)
Income	0.054	0.235**	0.015	0.016*
	(0.061)	(0.107)	(0.010)	(0.009)
Prior AI Exp	-0.022	-0.628***	-0.058***	-0.058***
	(0.091)	(0.161)	(0.013)	(0.014)
Ln (Duration)	0.078**	0.315***	0.013*	0.014**
	(0.030)	(0.087)	(0.007)	(0.007)
Independent Variables:				
Proactivity	-1.038***	0.337**	0.005	0.005
	(0.106)	(0.165)	(0.014)	(0.014)
Social identity	0.006	0.034	0.013	-0.009
	(0.086)	(0.157)	(0.014)	(0.016)
Inefficiency		-3.491***	-0.072**	-0.141***
		(0.431)	(0.034)	(0.051)
Inefficiency*Social identity				0.134**
				(0.066)
Constant	-2.237***	-1.124**	1.263***	1.269***
	(0.259)	(0.515)	(0.047)	(0.047)
Observations	1,020	1,020	1,020	1,020
(pseudo) Adjusted R-squared	0.053	0.142	0.029	0.033

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.10. Randomization Check

Group	Proactivity Design	Social identity Design	N	Age	Income	Gender	Prior AI Exp
Group1	Proactivity	Social identity	253	2.636	2.407	0.060	0.664
Group2	Proactivity	Non-social identity	257	2.630	2.428	0.160	0.654
Group3	Non-proactivity	Social identity	263	2.624	2.437	0.103	0.597
Group4	Non-proactivity	Non-social identity	247	2.696	2.328	0.093	0.636
F-value				0.410	1.080	4.800	0.980
P-value				0.748	0.355	0.003	0.403

Note1: Age is measured by age groups from 1 to 4.

2.7. Discussion and Implications

The research was conducted through an online field experiment. The research mainly focused on how chatbot designs influence communication inefficiency from a perspective of uncertainty reduction. Specifically, the study developed a measurement of human-chatbot communication inefficiency by identifying four types of communication inefficiency—misunderstanding between consumers and chatbots, conversation diverging on irrelevant topics, asking the other party to repeat or rephrase questions, and guessing the meaning of expressions. This randomized online experiment indicated that: first, the chatbot proactive design decrease communication inefficiency by 15.8%. Then, communication inefficiency negatively impacted consumer satisfaction and consumer task success. When one unite increases in communication inefficiency, the task success and consumer satisfaction decrease about 96% and 38% respectively. In addition, the endowment of social identity in human-chatbot interaction weakened the negative relationship between communication inefficiency and consumer satisfaction. This research provides foundations for future human-AI interaction research and gives some advice for practitioners.

For theoretical implications, the current study contributes to the literature in three ways. First, the research introduced the uncertainty reduction theory to the field of human-chatbot communication as the key theoretical foundation. In the existing literature, uncertainty reduction theory has been mainly applied in intercultural communication (i.e., Nelson, 1992), job hiring process research (i.e., Ragan, 1983), and computer-mediated communication research (i.e., Pratt et al., 1999; Antheunis et al., 2011). As articulated at the beginning of this study, uncertainty is a key obstacle to a smooth human-chatbot conversation. Thus, the uncertainty reduction theory fits well in the human-chatbot conversation context and can shed light on the process. In general, there are two types of uncertainty—cognitive uncertainty and behavioral uncertainty. This research adopted these two notions to the field of human-chatbot interaction and examines two features of chatbot design—the proactivity design for cognitive uncertainty reduction and the social identity design for behavioral uncertainty reduction. Furthermore, the impacts of these two designs have been examined on two types of human-chatbot conversation outcomes—task success and user satisfaction. We found that, in the context of human-chatbot interaction, cognitive uncertainty reduction serves as the main mechanism underlying the impacts of the proactivity design on the two outcomes.

Second and following the uncertainty reduction perspective, we conceptualized and operationalized communication inefficiency as a mediator between the two chatbot designs and consumer satisfaction and consumer task success. Most existing literature investigated how chatbot designs impact on communication outcomes (i.e., purchase (i.e., Luo et al., 2019), persuasion (i.e., Shi et al.,2020), subjective perception of consumers towards chatbots (i.e., Ashktorab et al., 2019; Cho, 2019; Tärning and Silvervarg, 2019)). There is a small body of research that went into the conversational process and studied how chatbot designs influence the human-chatbot communication inefficiency and, in turn, impact on task outcome and affective outcome. This study integrated two relevant notions—non-progress (Li et al., 2020) and obstacles (Myers et al., 2018) from the Human-Computer interaction (HCI) literature—as the conceptual foundation of communication inefficiency and developed a metric for this variable—the number of the occurrences of communication non-smoothness per turn in human-chatbot communication. We summarized four types of communication inefficiency in human-chatbot communication—(1) misunderstanding between consumers and chatbots, (2) conversation diverging on irrelevant topics, (3) asking the other party to repeat or rephrase questions, and (4) guessing the meaning of expressions—and found that communication inefficiency mediated the positive relationships between chatbot proactivity designs and the two outcomes—consumer task success and consumer satisfaction.

Finally, although the social identity design model does not have significant direct impacts on human-chatbot conversation outcomes, this research utilized social identity theory and investigated the moderation effect of social identity design on the relationship between human-chatbot conversational process and consumer satisfaction. Recently, more attention has focused on the direct effects of social identity cues in chatbot literature (i.e., Araujo, 2018; Schanke et al., 2021). Its role as a boundary condition is not thoroughly examined. This research examined the moderation effect of social identity design on the relationship between communication inefficiency and consumer satisfaction by assigning a community member name to the chatbot and using intimate language in communications. In accordance with social identity theory, individuals tend to treat ingroup members favorably (Taylor and Doria, 1981). Thus, the social identity design would have a moderation effect on the relationship between ineffective communication and consumer satisfaction. The empirical examination supported the hypothesis and revealed the social identity design weakens the negative relationship between communication inefficiency and consumer satisfaction.

In practice, this research gave us a better understanding of chatbot proactivity and social identity designs in a business setting, which plays a pivotal role in the commercial usage of AI applications. The experiment is conducted in a real business environment cooperating with a famous beer brand in southeast China, which is valuable for practitioners. The findings provided evidence for practitioners who plan to launch chatbots in their businesses. In particular, this study further studied the human-chatbot conversational process and provided a method of measuring human-chatbot communication performance by evaluating the communication inefficiency between humans and chatbots. Compared to existing human-chatbot communication measurement—counting turns, this measurement helps companies better understand the inefficiency of human-chatbot communications.

Second, the research also provided insights into the effects of the chatbot proactivity design and social identity design. The two designs are popular in the industry and widely discussed. This research provided empirical evidence of the two designs in a business setting and provided an underlying mechanism from the uncertainty reduction perspective. We studied the effects of chatbot designs on the human-chatbot communication inefficiency and the moderation role of social identity on the relationship between communication inefficiency and customer satisfaction. As our findings revealed, chatbot proactivity design is a beneficial strategy for companies to reduce human-chatbot communication inefficiency. The social identity design can help weaken the negative effect of human-chatbot communication inefficiency on customer satisfaction. Especially, the type of social identity given to the chatbot needs to be carefully considered by practitioners.

Third, the two dependent variables—task success and consumer satisfaction—are essential in business. In the industry, conversion and the user experience are the critical metrics of chatbot performance. The two dependent variables are highly related to the two chatbot performance metrics, respectively. Especially, the chatbot is the first contact with potential consumers. Human-chatbot communication task success and consumer satisfaction help to understand the chatbot performance. High customer satisfaction and a successful conversation help companies enhance relationships with customers and, in the end, earn customers' loyalty.

2.8. Limitations

This research is subject to some limitations. First, the experiment was conducted in the context of the beer industry, and, therefore, the findings may not be applicable to other industries. Second, this research focused on task-oriented human-chatbot communication and, thus, the findings might be different in non-task-oriented human-chatbot communication. Third, chatbot proactivity design and social identity design may lead to other consequences which were not considered in this research.

Chapter 3 : User Evaluation of Human-AI Hybrid Service

Abstract

The study investigates how the introduction of a hybrid human-AI service affects user evaluations of the service. AI and humans do have their own advantages when facing different tasks in digital platforms—for example, AI agents are better skilled in reliability, scalability, speed, accuracy, and generalization, while human agents are better skilled in creativity, judgment, and empathy (Rai and Sarker, 2019). A Human-AI hybrid service system can benefit from both AI and Human agents' advantages, and it is of interest to study the impact of the hybrid services in digital platforms. In particular, our study examines the associations between the human-AI hybrid services and user evaluations of the service in the context of a major audio streaming platform. On the digital platform, the introduction of AI podcasters creates the phenomenon—for the same audio, sometimes the AI voice version is released first, and sometimes the human voice version is released first. This provides us with a great opportunity to study how the introduction of a human-AI hybrid service affects user evaluations of the service from a temporal perspective. The empirical findings show that the user evaluations of those who are first exposed to human service are improved with the presence of the human-AI hybrid service, and the user evaluations of those who are first exposed to AI service get worse with the presence of the human-AI hybrid service. Our findings have important implications for both digital platforms and AI developers.

3.1. Introduction

Artificial Intelligence (AI) has been transforming the world in many ways, such as personalized shopping, fraud prevention in E-commerce, autonomous vehicles in transportation, removing propaganda content, and engaging users on social media[7]. Though recent development of the technology allows AI agents to have reasoning, problem-solving,

[7] https://www.simplilearn.com/tutorials/artificial-intelligence-tutorial/artificial-intelligence-applications

learning, and natural language processing abilities, AI agents still cannot fully replace the human agents (Lee, 2021) since AI and humans do have their own advantages when facing different tasks—while some features of digital platform tasks match well with AI agent skills (e.g., reliability, scalability, speed, accuracy, generalization), others are likely to correspond better with human agents (e.g., creativity, judgment, empathy) (Rai and Sarker, 2019). In the case of online sales assistants, AI agents can well handle repetitive and frequently asked questions and human agents can better handle more specific and flexible questions from clients. Many websites provide customers with both human and AI sales assistants. Customers can choose the sales assistant service according to their preferences. It is becoming increasingly common to offer both AI and human services: e.g. online language learning, after-sales service, pre-sales customer service. A recent survey shows that the collaboration of human and AI agents achieves the best performance by enhancing each other's complementary strengths[8].

The collaboration of human and AI agents refers to the broad involvement of both human and AI agents in the completion of a task (Lai et al., 2021), which can be AI replacing humans, humans and AI augmenting each other, or AI and humans are assembled (Rai and Sarker, 2019). Existing literature also documents similar constructs including Human-AI collaboration (Lai et al., 2021; Wang et al., 2020), Human-AI hybrids (Pereira et al., 2021; Rai and Sarker, 2019), Human-AI hybrid intelligence (Dellermann et al., 2019), and others. Based on the human-AI hybrids construct (Rai and Sarker, 2019), our study focuses on the human-AI hybrid service, which is defined as the human and AI agents providing service to users on the *same* task. We study the effect of introduction the human-AI hybrid service on an online streaming platform. The introduction of the Human-AI hybrid service in the online streaming digital platform Ximalaya (see Figure 3.1), creates a new phenomenon—the platforms can only offer human services first in the past (Human service → Human-AI hybrid services), but now can offer AI

[8] https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces

services before human services (AI service → Human-AI hybrid services) because of the speed and generalization characteristics of AI. The platform provides listeners with news audios, book audios, and talk show audios and acts as an intermediary between podcasters and listeners. For example, users can listen to news, fictions, or talk shows on the platform while waiting for the bus or doing housework. It helps users make the best use of their time to access information. Similar APPs include Audible, Audiobooks, and so on. This platform was founded in 2012 and had 250 million active monthly users in 2021[9]. At the end of 2020, the platform introduced AI podcasters. Human voices have the advantage of correct emotion and less mistakes, while AI voice will have the advantage of efficiency (i.e., upload 10 audios a day). For AI audios, they display AI voice icon in the page.

The introduction of AI podcasters creates the phenomenon—for the same audio, sometimes the AI voice version is released first, and sometimes the human voice version is released first. This provides us with a great opportunity to study how the introduction of a human-AI hybrid service affects user evaluations of the service. Given the different versions of services provided at first, we collect the review of the first version service and examine how user evaluations change after the later introduction of the human-AI hybrid service (see Figure 3.2).

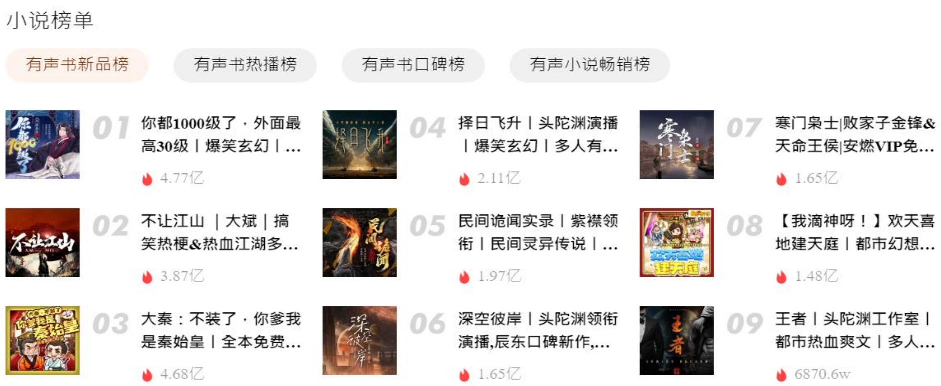

Figure 3.1. the Screenshot of the Platform

9 https://www.scmp.com/business/banking-finance/article/3148241/podcasting-platform-ximalaya-shelves-planned-us-ipo

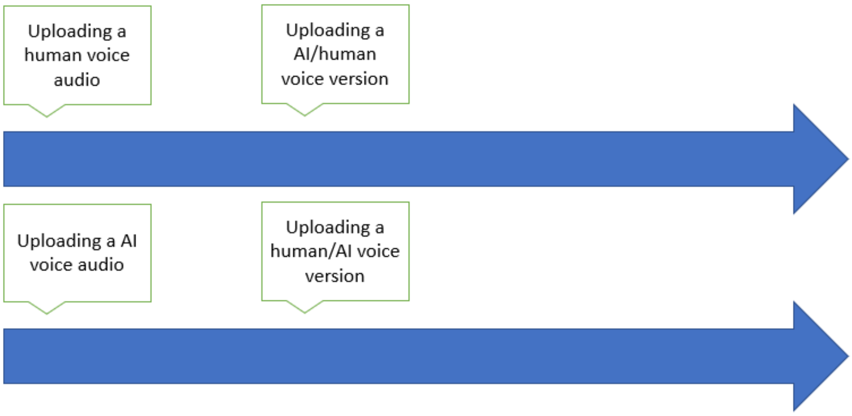

Figure 3.2. a Simulation of the Study Context

To the best of my knowledge, little research has been done on this new phenomenon and it remains unknown about user evaluations with the presence of the human-AI hybrid service in the digital platform. On the one hand, the user evaluations are possibly more affected by the first service that they received (the first mover advantage), although the platform later offers other options. On the other hand, the user evaluations are possibly driven by the contrast between AI service and human service (the contrast effect). Both ways are theoretically conceivable, providing a good opportunity to empirically examine the two competing effects. In addition, the effect of introducing a human-AI hybrid service on user evaluations of the service varies according to different service content characteristics (i.e., the popularity). For example, when the audio is broadcasted by a superstar, the effect of introduction human-AI hybrid service may be weakened, and people's attention may be diverted by the popularity. Given these, we propose the following research questions:

1. How does the introduction of the human-AI hybrid service affect user evaluations of the original service?
2. How do the effects of introducing a human-AI hybrid service on user evaluations vary with different service content characteristics?

Utilizing the data from a major online audio streaming platform, we would like to explore the associations between the human-AI hybrid service and user evaluations of the service. We investigate two situations of the human-AI hybrid service implementations — 1) the user is first exposed to an AI service, 2) the user is first exposed to a human service. Then the platform provides a different version of service on the same task (the implementation of a human-AI hybrid service). The analysis shows that if the user was first exposed to an AI service, the later introduction of a human-AI hybrid service is negatively associated with the average user evaluation of the service. If the user was first exposed to a human service, the later introduction of a human-AI hybrid service is positively associated with the average user evaluation of the service. Our findings are consistent with the views that the users generally prefer human services (Longoni et al., 2019, Luo et al., 2019, Mou and Xu, 2017), and confirm the contrast effect story. Moreover, we find that the effect is more pronounced in low-popularity conditions such as audios with a fewer-play count, a lower-podcaster level, and a fewer-number of characteristics.

This study intents to make contributions to the Information System literature in three aspects. First, we investigate the interesting phenomena in which both human and AI provide service on the same task. We examine this problem from a temporal perspective rather than a static analysis. Second, in academia, most literature studied this topic on two streams—human-AI collaboration system design (e.g., Correia et al., 2020; Ostheimer et al., 2021) and compare the human-only mode, the AI-only mode, and the human-AI coexistence mode (e.g., Peeters et al., 2021; Sowa et al., 2021). The user evaluation on the service of the human-AI hybrid service is still not thoroughly understood. Therefore, there is a need to advance our understanding of the user evaluation on the human-AI hybrid service. This research contributes to the literature by studying the human-AI hybrid service from the users' side. Third, this research contributes to AI in digital platforms literature and AI introduction in business literature. Rich existing literature studied the introduction of AI in digital platforms (Alt, 2021; Colace et al., 2018). They mainly studied how AI supports digital platform processes (Alt, 2021; Mucha and

Seppälä, 2020) and AI as a service on the digital platforms (Alt, 2021; Colace et al., 2018). In AI in business research, scholars mainly discussed the antecedents (i.e., Baabdullah et al., 2021; Ostrom et al., 2019) and consequences (i.e., Luo et, al., 2019; Schuetzler et al., 2020; Shi et al., 2020) of introduction AI applications. Little research studied human and AI hybrid services as a whole, although the human-AI hybrid service is gaining momentum. This study will contribute to the literature by filling the research gaps.

The paper is organized as follows. In section two, existing literature will be reviewed. Next, hypothesis development will be introduced. Then, an empirical examination will be provided, which includes data and variables, analysis results, robustness check, and subsample analysis. The final section will be the conclusion and discussion.

3.2. Literature Review

Given the research objectives identified, a review of the relevant literature on human-AI hybrid service is conducted in Information Systems, Computer Science, and Management fields. We review the literature from three aspects: (1) human-AI collaboration literature, (2) AI in digital platform literature, and (3) AI introduction in business literature.

3.2.1. Human-AI Collaboration

In order to clearly conceptualize human-AI hybrid service. We first identified some similar human-AI hybrid service constructs in Table 3.1. Based on human-AI hybrids construct (Rai and Sarker, 2019), the human-AI hybrid service is defined as the human and AI agents providing service to users on the same task. Then, as human-AI collaboration is the broadest construct, we review related literature on the human-AI collaboration.

Table 3.1. Human-AI Collaboration Related Constructs

Constructs	Description	Focus	How AI involved
Human-AI collaboration	"The collaboration between single or multiple humans and AI systems. …It is an evolving, interactive process whereby two or more parties (human and AI) actively and reciprocally engage in joint activities aimed at achieving one or more shared goals" (Lai et al., 2021, p390).	Humans and AI work together to achieve a goal. It is a broad construct that encompasses several of the following constructs, such as human-AI hybrids, human-AI hybrid intelligence, and human-AI hybrid service.	Complete part of the task or the whole task.
Human-AI hybrids	There are several forms of human-artificial intelligence hybrids： (1) human-AI interdependence from substitution (AI replaces humans) (Rai and Sarker, 2019). (2) augmentation (humans and AI augment each other) (Rai and Sarker, 2019).	Humans and AI are working as a whole. For example, AI replaces humans in some positions.	Complete part of the task or the whole task.

	(3) assemblage (AI and humans are dynamically brought together to function as an integrated unit) (Rai and Sarker, 2019).		
Human-AI hybrid intelligence	Using the complementary strengths of human intelligence and AI so that they can perform better than each of the two could separately (Dellermann et al., 2019).	It focuses on optimizing performance by integrating humans and AI using the complementary strengths	Complete part of the task.
Human-AI hybrid service	It is extended from the human-AI hybrid construct (Rai and Sarker, 2019). Humans and AI provide services to users on the same task on digital platforms. Users will choose the service according to their preferences.	Two services (human and AI) exist simultaneously for the same task. For example, platforms provide both human sales assistants and AI sales assistants to users.	Complete the whole task.

We recognize two streams of human-AI collaboration literature. One stream of literature is on human-AI collaboration system design. The second stream is on the comparison among the human-only mode, the AI-only mode, and the human-AI coexistence.

One stream of research on the human-AI coexistence mainly focuses on discussing human-AI collaboration systems design principles, such as understanding the characteristics and limits of AI (Korteling et al., 2021), understanding the biases in AI and human cognition

(Korteling et al., 2021), humans should control AI and trust AI (Ostheimer et al., 2021), and explainable human-AI system principles (Mueller et al., 2020). Some studies also proposed the specific human-AI coexistence workflow and taxonomy of the human-AI coexistence system (Correia et al., 2020; Dellermann et al., 2021; Dubey et al., 2020). Notably, some studies further discuss how to apply human-AI hybrids in various contexts. For instance, the workflow of human-AI coexistence was developed to review literature in academia (Thomas et al., 2017). A conceptual framework for human-AI hybrid adaptivity was proposed in an education context (Holstein Kenneth and Aleven, 2020). Another model of human-AI was presented to create songs (Huang et al., 2020). Although human-AI coexistence is already quite popular on digital platforms, human-AI coexistence designs are only discussed in limited contexts.

The other stream of studies on human-AI coexistence mainly discussed the performances of the human-only mode, the AI-only mode, and the human-AI coexistence. They examined the three modes with qualitative methods and quantitative methods (e.g., Bansal et al., 2019a; Sowa et al., 2021). Most studies suggest that human-AI collaboration has the best performance (e.g., Gao and Jiang, 2021; Rai and Sarker, 2019; Sowa et al., 2021). For example, compared with the human-only system, the human-AI coexistence system reduced response time, keystrokes were reduced, and more suggestions were adopted in the context of the human-chatbot conversation (Gao and Jiang, 2021). Human-AI collaboration also can increase productivity (Sowa et al., 2021). Further, some studies considered the performance of the human-AI coexistence from a dynamic perspective by examining the effects of changing AI capability, human cognition, or constraint conditions in human-AI coexistence. For instance, increased AI capability hurts human-AI collaboration performance (Bansal et al., 2019b). The trade-offs among accuracy, cost, and efficiency were also discussed in human-AI coexistence (Kahn et al., 2020). Humans' mental models influenced by AI's error boundary and task property is also positively related to human-AI coexistence performance (Bansal et al., 2019a).

Low human expectations toward AI lead to better evaluations of human-AI coexistence (Khadpe et al., 2020). Most studies from the above literature researched human-AI coexistence from the human-AI system performance perspective. However, little literature investigated human-AI coexistence from customers' attitudes.

3.2.2. AI on Digital Platform Literature

Existing literature studied the introduction of AI in digital platforms (Alt, 2021; Colace et al., 2018). They mainly studied how AI supports digital platform processes (i.e., Alt, 2021; Mucha and Seppälä, 2020) and AI as a service in digital platforms (i.e., Alt, 2021; Colace et al., 2018). AI mainly support digital platform process such as recommendation systems (Malgonde et al., 2020) and process automation systems (Hofmann et al., 2020). AI can improve the performance of digital platforms by improving efficiency, accuracy, and diversity. For example, Microsoft implemented automatic spam filtering on its platform in 2006, and eBay used AI to enhance product categorizations and product searches (Mucha and Seppälä, 2020). AI also can be used to reduce bias (such as gender bias and developing country bias) on labor platforms (Rai and Sarker, 2019). In addition, the role of AI as a service is also researched in the existing literature in digital platform literature (e.g., Illescas-Manzano et al., 2021; Vanichvasin, 2021). AI is considered a useful service in digital platforms (i.e., Brynjolfsson, 2019; Vanichvasin, 2021). For instance, the chatbot is a typical AI service on digital platforms. Chatbots can help knowledge sharing in education platforms by increasing fun and efficiency (Vanichvasin, 2021). Another example is AI as a translation tool on the digital platforms. The AI translation service increases exports by 10.9% in e-commerce platforms by reducing the language barrier (Brynjolfsson, 2019). Despite the fact that most researchers believe the introduction of AI benefits digital platforms (i.e., Brynjolfsson, 2019; Rai and Sarker, 2019; Vanichvasin, 2021), most research ignores the fact that AI cannot completely replace humans. Rather than concentrating solely on the introduction of AI, we should take both human and AI services into consideration. The human-AI hybrid service is an area that is still under-researched, particularly from a temporal perspective. Therefore, further research on human-AI hybrid service is highly valued.

3.2.3. AI introduction in business literature

In business research, many scholars discussed the introduction of AI applications in business, which including the antecedents (i.e., Baabdullah et al., 2021; Holmström, 2022; Jöhnk et al., 2021; Ostrom et al., 2019) and consequences (i.e., Chung et al., 2020; Gill, 2020; Luo et, al., 2019; Schuetzler et al., 2020; Shi et al., 2020) of introduction AI applications. Compared to AI consequences research, AI antecedent research is still at an early stage. In AI antecedent research, the AI readiness construct is widely discussed (i.e., Baabdullah et al., 2021; Holmström, 2022; Jöhnk et al., 2021) and defined as “an organization’s abilities to deploy and use AI in ways that add value to the organization” (Holmström, 2022, p330). Strategic alignment, resources, knowledge, culture, and data are five factors of AI readiness (Jöhnk et al., 2021). Two theories are mainly discussed as underlying mechanisms in AI readiness research—the technology organization environment (TOE) and Diffusion of Innovation (DOI) (Alsheibani, 2018; Najdawi, 2020). AI readiness is mainly studied in the context of organizations (i.e., Alsheibani et al., 2018; Frick, 2021; Jöhnk et al., 2021) and education (i.e., Karaca et al., 2021; Luckin et al., 2022). For example, a non-significant relationship was found between empowering leadership and AI readiness (Frick, 2021). In AI consequences research, many business consequences are discussed including purchase (i.e., Luo et al., 2019; Schanke et al., 2021), donation persuasion (i.e., Shi et al., 2020), and brand engagement (i.e., Kull et al., 2021). For example, Luo’s research revealed the disclosure of AI identity will decrease purchases by almost 80% (Luo et al., 2019). In the workplace, the introduction of AI has a negative influence on employees' identification with their occupations and triggers people's anxieties about being replaced (Mirbabaie et al., 2022). AI applications have already been applied in various situations such as the workplace (i.e., Mirbabaie et al., 2022; Yu, 2019), sales (i.e., Luo et al., 2019; Schanke et al., 2021), and education (i.e., Boulay, 2016; Kim et al., 2020). We are already living in a world where humans and AI coexist. In many scenarios, users are given options for both AI services and human services. In academia, we cannot just discuss the influences of AI introduction; we have to think about a society where humans and AI work together to provide services to consumers. Given the business and theoretical significance of human-AI hybrid services, the need for a better understanding of human-AI hybrid service's effects on the individuals' user evaluation on the service is growing. Therefore, we identified the following research gaps: (1) Most studies from the above literature review researched human-AI coexistence from the human-AI system performance perspective. However, little literature investigated human-AI coexistence from customers' attitudes. (2) We are already living in a world where humans and AI coexist. In many scenarios, users are given options for both AI services and human services. In academia, we cannot just discuss the impact of introducing AI. We have to think about a world where humans and AI work together to provide services to consumers. (3) Human-AI hybrid service is still under-researched, particularly from a temporal perspective.

3.3. Hypothesis Development

This study will apply the contrast effect as the underlying mechanism to develop the hypothesis. We choose to build a hypothesis using the contrast effect instead of the first mover advantage because the performance of AI in our research context falls short of user expectations. Moreover, users perceive human service providers and AI service providers differently, leading them to naturally compare the two (Chen, et al., 2021). This indicates that the contrast effect is likely to have a stronger influence on user perceptions and preferences compared to the first mover advantage, which typically applies to very similar products like Coca-Cola and Pepsi. The contrast effect was first noted in the 17th century by observing the temperature of lukewarm water depending on whether the hand has previously been in hot or cold water (Kushner, 2008). After that, many other domains also have revealed this phenomenon, such as personality trait judgment (e.g., Herr, 1986; Manis et al., 1988) and sensory perception (e.g., Sarris, 1967; Pol et al., 1998). According to Hovland, Harvey, and Sherif (1957), the contrast effect in social judgment refers to a person's propensity to exaggerate the gap between two views. It is an unconscious bias that results from evaluating two items in comparison rather than separately.

The average user evaluation of the service is used as the dependent variable, which is a subjective feeling. In this case, users generally consider human service to be better (Longoni et al., 2019, Luo et al., 2019, Mou and Xu, 2017). Based on the contrast effect, if the AI service were to be evaluated on its own, users would give it a score of 7. With human service as a comparison, the rating for AI voice may drop to 6. Conversely, a human service would be given a rating of 8 but could potentially rise to 9 in comparison to the AI service.

If users receive AI services first, users have worse user evaluations of the service after the later introduction of a human-AI hybrid service. If users receive human services first, users have better user evaluations of the service after the later introduction of a human-AI hybrid service. Users generally consider human service to be better (Longoni et al., 2019, Luo et al., 2019, Mou and Xu, 2017). For example, when customers are informed that they will communicate with a chatbot rather than a human agent, they encounter more uncertainty, anticipate a less positive experience, and have lower expectations in terms of their social presence (Edwards et al., 2016; Spence et al., 2014). After the introduction of other types of services with the same functions, users will compare them to the service they received. At that time, the contrast effect will enhance the differences (Hovland et al., 1957). In the situation of the online streaming platform, AI and human broadcasting services have their advantages, respectively. The human service has a more correct rhythm with fewer mispronunciations, while the AI service is more efficient (i.e., post ten audios in a day). After the contrast, the advantages of human service (i.e., fewer mispronunciations and right rhythm) are amplified, as are the disadvantages of AI service (i.e., mispronunciation). Therefore, in the introduction of human-AI hybrid service process:

H1: For users who first receive the human service, the later introduction of a human-AI service will positively impact on user evaluations of the service.

H2: For users who first receive the AI service, the later introduction of a human-AI service will negatively impact on user evaluations of the service.

The less popular the audio is, the more pronounced this effect will be. If an audio is very popular (e.g., high play count, broadcasted by a high-level podcaster, the audio has many characteristics), users' attention will be diverted by these. For example, popular audios will divert the listener's attention from the contrast between the human voice and the AI voice. When people's attention is distracted by other stimuli (e.g., audio characteristics, podcaster level), people will allocate less attentional resources to the original point (the contrast between human and AI voices) and, thus, process the information less efficiently (Johnson and Proctor, 2003). In the online streaming context, if the audio is very popular, the listeners' attention on the human-AI hybrid service is reduced. Therefore, if the audio is less popular, listeners will put more attention on the human-AI hybrid service contrast. Thus, the contrast effect is more pronounced.

H3: For users who first receive the human service,, the effect is more pronounced in less popular conditions.

H4: For users who first receive the AI service,, the effect is more pronounced in less popular conditions.

3.4. Data and Variables

As the introduction part introduced, the context of the research is on an online audio platform (see Figure 3.3). The platform introduced AI podcasters at the end of 2020, which creates the phenomenon—for the same audio, sometimes the AI voice version is released first, and sometimes the human voice version is released first. In this study, we examine how user evaluation on the service changes in the process of a human-AI hybrid service introduction, while considering the human-AI hybrid service release strategies (human voice→human-human voice vs. human voice→human-AI voice and AI voice->AI-AI voice vs. AI voice→human-AI voice) (see Figure 3.2). We only collected the reviews from the first launched version and checked how the average user evaluation on the service changed in the process of introducing a human-AI hybrid service. If there is a new version, it will appear on the page of the old version that is being listen to.

The sample consists of 334 audios. The treated group consists of the audios that provide the human-AI hybrid service. The control group consists of the audios only provide one type of service (AI voice service->AI-AI voice service or human voice service->human-human voice service). To match with the treated audio, the audios in the control group were selected following the matching steps (Loughran and Ritter, 1995): Step 1: Audios were ranked by their review NO. per month. The audio with all the same topics, all the same characteristics, the same podcaster NO., and the closest review NO. per month was selected. Step 2: if no audio was matched in step 1, The audio with all the same topics, all the same characteristics, and the closest review NO. per month was selected. Step 3: if no audio was matched in step 2, The audio with at least 50% same topics, at least 50% same characteristics, and the closest review NO. per month was selected. Finally, we selected 92 audios for human->hybrid type (treated vs. control: 46 vs. 46) and 242 audios for AI->hybrid type (treated vs. control:121 vs. 121).

In this study, the review positivity is used as the dependent variable representing the level of positivity of user evaluations of the service ranging from 0 to 1. The accuracy level is the level of accuracy in this Natural Language Processing (NLP) analysis ranging from 0 to 1. It is an assessment of this NLP analysis for the dependent variable review positivity. The positivity and the NLP accuracy level were generated by Baidu natural language processing API (https://cloud.baidu.com/doc/NLP/index.html), which is one of the most famous Chinese natural language processing services. Table 3.2 summarizes the variable definitions. Figure 3.3 displays how to extract the variables from the platform.

Figure 3.3. the Screenshot of the Platform

Table 3.2. Variable Definitions

Variables	Definitions
Review Positivity	the possibility of the review's emotion is positive.
After	whether the review was posted after the introduction of the second voice service.
Hybrid Service	whether the audio has both human and AI voice service.
NLP Accuracy	the level of accuracy in this natural language processing analysis (review positivity) ranging from 0 to 1.
Review NO.	the log of the number of reviews the audio has over one month.

Single Podcaster	whether the audio was podcasted by a single person.
Topics	the audio topics which were extracted from the audio tags.
Characteristics	the audio's characteristics which were generated by algorithms.

Table 3.3 provides the summary statistics for the two human-AI hybrid services forming situations. The variable *After* refers to whether the review was posted after the introduction of the hybrid voice service (yes:1 vs. no:0). The average *review positivity* represents the average user evaluation on the service, and the mean is 0.493 and 0.299 respectively, which is consistent with prior research—customers prefer human service than AI service (Luo et al., 2019). *Hybrid service* refers to whether the second version of the audio is broadcasted by another type of podcaster (yes:1 vs. no:0)—if the original version was broadcasted by a human, the second voice is AI voice or vice versa (mean=0.507 and 0.679). *Ln(review NO.)* refers to the log of the number of reviews the audio has over one month. The averages are 1.302 and 0.968 in the two situations, with a standard deviation of 0.587 and 0.425. *Ln(NLP accuracy)* refers to the average accuracy of NLP analysis. The averages are -0.223 and -0.194 in the two situations. *Single Podcaster* refers to whether the audio is podcasted by a single person (yes vs no: 1vs 0). It reveals that most of the audios are podcasted by a single person in the two situations (mean=0.835 and 0.537). The four *topics* represent the audio content categories including righteous ardor topic(热血), reborn topic(重生）, fantasy topic(奇幻）, and modern topic(现代）. They were extracted from the audio tags. The last four characteristics were extracted by the TextRank algorithm, which is an extractive summarization technique (Balcerzak et al., 2014). We first extract the text from the corresponding book. Next, we apply the Textranking algorithm to identify key features within the text. These features are then clustered using Support Vector Machines (SVM) to categorize them as distinct characteristics. The

characteristics represent the content features of the audios. They are romance, Kungfu, food, and other characteristics. Table 3.4 displays the correlations among the variables.

Table 3.3. Summary Statistics

Panel A: Summary Statistics of Human->Hybrid			
VarName	Mean	SD	Median
Review Positivity	0.493	0.343	0.495
After	0.493	0.501	0.000
Hybrid Service	0.507	0.501	1.000
Ln(NLP accuracy)	-0.223	0.507	-0.076
Ln(Review NO.)	1.302	0.587	1.099
Single Podcaster	0.835	0.372	1.000
Topic_righteous ardor	0.020	0.140	0.000
Topic_reborn	0.072	0.259	0.000
Topic_fantasy	0.749	0.434	1.000
Topic_modern	0.085	0.280	0.000
Characteristic_romance	0.088	0.284	0.000
Characteristic_kungfu	0.009	0.092	0.000
Characteristic_food	0.046	0.209	0.000
Characteristic_other	0.630	0.484	1.000
Panel B: Summary Statistics of AI->Hybrid			
VarName	Mean	SD	Median
Review Positivity	0.299	0.365	0.088
After	0.365	0.482	0.000
Hybrid Service	0.679	0.467	1.000
Ln(NLP accuracy)	-0.194	0.423	-0.035
Ln(Review NO.)	0.968	0.425	0.693
Single Podcaster	0.537	0.499	1.000
Topic_righteous ardor	0.000	0.000	0.000
Topic_reborn	0.159	0.366	0.000
Topic_fantasy	0.019	0.137	0.000
Topic_modern	0.696	0.460	1.000
Characteristic_romance	0.074	0.263	0.000
Characteristic_kongfu	0.042	0.202	0.000
Characteristic_food	0.102	0.303	0.000
Characteristic_other	0.495	0.501	0.000

Table 3.4. Correlation Table

Correlation Coefficient of Human→Hybrid						
	(1)	(2)	(3)	(4)	(5)	(6)
Review Positivity	1.000	0.038	-0.432	-0.038	0.025	-0.249
After	0.027	1.000	-0.187	0.024	-0.135	0.371
Hybrid Service	-0.398	-0.179	1.000	0.003	0.104	0.438
Ln (NLP accuracy)	-0.011	-0.033	0.046	1.000	-0.263	-0.073
Ln (Review No)	0.004	-0.122	0.117	0.118	1.000	0.093
Single Podcaster	-0.241	0.377	0.451	-0.005	0.120	1.000
Correlation Coefficient of AI→Hybrid						
	(1)	(2)	(3)	(4)	(5)	(6)
Review Positivity	1.000	-0.042	-0.169	-0.532	0.127	0.160
After	-0.022	1.000	-0.088	0.082	-0.254	0.066
Hybrid Service	-0.261	-0.074	1.000	-0.058	0.285	0.050
Ln (NLP accuracy)	-0.202	0.024	-0.019	1.000	-0.295	-0.167
Ln (Review No)	-0.066	-0.246	0.277	0.038	1.000	0.283
Single Podcaster	0.027	0.058	0.038	0.022	0.282	1.000

Note: Lower-triangular cells report Pearson's correlation coefficients, upper-triangular cells are Spearman’s rank correlation

This model is used to examine the hypothesis. In the model, we have control groups (AI service→AI-AI service and human service→human-human service) and treated groups (AI service→human-AI hybrid service and Human service→human-AI hybrid service). The model is developed as follows:

$$Avg\ user\ evaluation_{it} = \beta_0 + \beta_1 Hybrid\ service_i + \beta_2 After + \beta_3 Hybrid\ service_i * After + \beta_4 Z_{it} + \epsilon_{it}$$

In the model, i refers to an audio. t refers to time. Where the average user evaluation of the service is the average review positivity of audio i over month t, Hybrid service is whether the audio i provide both human and AI broadcasting service. Z is the vector of remaining control variables, including the review NO. of the audio i at month t, podcaster NO. (single: yes:1, no:0) of audio i, the audio topics of the audio i, NLP accuracy of these reviews, and the audio characteristics of audio i. *After* refers to whether the review was posted after the introduction of the human-AI hybrid service.

3.5. Results

We estimate the main effect of human-AI hybrid service on the average user evaluation on the service using the model. We analyze the model in the time period of T-10 vs. T+10 (see Table 3.5). T-10 and T+10 refer to the time periods of the first ten months before or after the introduction of the human-AI hybrid service. The control group is non-treated audios that do not provide human-AI hybrid service. Both the Linear Probability Model (LPM) and the Tobit model are employed, with the latter serving as a robustness test for the former. The linear probability model aids in the clear interpretation of coefficients (Sun et al., 2021), and the review positivity in our regression fall within the [0, 1] range. Alternatively, Tobit model is appropriate for censored dependent variables.

In this study, two hybrid service forming strategies are considered. The model compares the human voice→the human-human voice vs. the human voice→the human-AI voice (see Table 3.5 columns 1 and 2). The other model compares the AI voice→the AI-human hybrid voice vs. the AI voice→the AI-AI voice (see Table 3.5 columns 3 and 4). The results reveal that if users are receiving human services, the introduction of a human-AI hybrid service is associated with a 34.9% user evaluation growth (see Table 3.5 column 1), supporting H1. If users are receiving AI services, the introduction of a human-AI hybrid service is associated with a 19.8% drop in user evaluation on the service (See Table 3.5 column 3), supporting H2.

Table 3.5. Main Results

Dependent Variable:	Review Positivity			
Hybrid Status:	Human->Hybrid		AI->Hybrid	
Model:	LPM	Tobit	LPM	Tobit
Hybrid Service*After	**0.349****	**0.349*****	**-0.198***	**-0.198****
	(2.080)	**(2.622)**	**(-1.927)**	**(-2.302)**
After	0.001	0.001	0.159	0.159*
	(0.010)	(0.012)	(1.595)	(1.937)
Hybrid Service	-0.679***	-0.679***	-0.237***	-0.237***
	(-3.518)	(-4.242)	(-3.131)	(-3.550)
Ln (NLP Accuracy)	-0.009	-0.009	-0.188***	-0.188***
	(-0.336)	(-0.253)	(-5.330)	(-4.998)
Ln (Review NO.)	0.059**	0.059*	0.005	0.005
	(2.038)	(1.926)	(0.127)	(0.109)
Single Podcaster	-0.002	-0.002	-0.021	-0.021
	(-0.025)	(-0.029)	(-0.400)	(-0.441)
Topics	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Constant	0.999***	0.999***	0.424***	0.424***
	(5.618)	(6.057)	(4.888)	(5.192)
N	342	342	464	464
Adjusted r^2 /Pseudo r^2	0.204	0.404	0.124	0.201

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.6. Robustness Check

The analysis reveals the positive relationship of introduction a human-AI hybrid service in the human service first situation and the negative relationship of introduction a human-AI hybrid service in the AI service first situation. To further validate the robustness of the results, we employ the placebo test to validate the results by artificially advancing the time of the second voice introduction to one month and two months respectively (Table 3.6). Table 3.6 reveals that if the *after* advances 1 or 2 months, the results are not significant. Both robustness checks are consistent with our main findings. In addition, we control play counts and ratings of audios in the model (Table 3.7). They reveals that our results are robust to multiple robustness checks. Therefore, we conclude the finding that the introduction of the human-AI hybrid service impacts user evaluations of the service. The results reveal that if users are receiving human services first, the later introduction of human-AI hybrid service is positively associated with user evaluations. If users are receiving AI services first, the later introduction of human-AI hybrid service is negatively associated with user evaluations. Overall, the introduction of human-AI hybrid service made user evaluations more polarised for existing users.

Table 3.6. Robustness Check

Dependent Variable:	Review Positivity			
Hybrid Status:	Human->Hybrid		AI->Hybrid	
Placebo test:	1 month	2 months	1 month	2 months
Hybrid Service*After	**0.201**	**0.268**	**-0.091**	**-0.064**
	(1.001)	**(1.574)**	**(-0.939)**	**(-0.634)**
After	-0.085	-0.199	0.115	0.042
	(-0.451)	(-1.240)	(1.193)	(0.388)
Hybrid Service	-0.480**	-0.486***	-0.270***	-0.246***
	(-2.431)	(-3.037)	(-3.595)	(-3.079)
Ln (NLP Accuracy)	-0.002	-0.009	-0.183***	-0.182***
	(-0.076)	(-0.338)	(-5.295)	(-5.199)
Ln (Review No.)	0.054*	0.051*	0.019	0.018
	(1.853)	(1.759)	(0.566)	(0.543)
Single Podcaster	0.139	0.113	-0.041	-0.031
	(0.692)	(0.683)	(-0.822)	(-0.609)
Topics	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Constant	0.777***	0.754***	0.440***	0.413***
	(6.081)	(6.384)	(4.918)	(4.386)
N	342	342	464	464
Adjusted r2	0.191	0.204	0.124	0.115

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.7. Robustness Check with Play Count and Ratings

	(1) Positive Probability	(2) Positive Probability	(3) Positive Probability	(4) Positive Probability
Hybrid Service*After	0.358**	0.358***	-0.202*	-0.202**
	(2.159)	(2.625)	(-1.939)	(-2.329)
Hybrid Service	-0.534**	-0.534***	-0.232	-0.232
	(-2.517)	(-2.774)	(-1.102)	(-0.582)
After	0.000	0.000	0.160	0.160*
	(0.002)	(0.002)	(1.593)	(1.946)
Play count	0.003**	0.003*	0.002	0.002
	(2.291)	(1.734)	(0.353)	(0.428)
Ln(Review NO)	0.023	0.023	0.001	0.001
	(0.607)	(0.602)	(0.017)	(0.015)
Ln(Rating)	1.296	1.296	-0.054	-0.054
	(1.458)	(1.525)	(-0.077)	(-0.026)
Ln(Confidence)	-0.006	-0.006	-0.188***	-0.188***
	(-0.231)	(-0.171)	(-5.298)	(-4.986)
Single Podcaster	-0.006	-0.006	-0.022	-0.022
	(-0.059)	(-0.064)	(-0.168)	(-0.080)
Topics	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time Fixed Effect	Yes	Yes	Yes	Yes
Constant	-1.964	-1.964	0.542	0.542
	(-0.961)	(-1.007)	(0.317)	(0.112)
N	342	342	464	464
Adjusted(pseudo) r^2	0.208	0.418	0.120	0.201

t statistics in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.7. Subsample Analysis

We further explore the performance among various audio features (H3 and H4) by subsample analysis. We classify the popularity based on the polycount median. First, we compare results in fewer play count audios and more play count audios (Table 3.8). Play count represents the popularity level of the audio. We find if the audio is played less often, the results are more pronounced in both situations (support H3 and H4). In fewer play count situations, the users who are first exposed to human voice increase the average user evaluation by 50.4% (see Table 3.8 column1). The users who are first exposed to AI voice decrease the average user evaluation by 21.7% (see Table 3.8 column 3). In addition, it is possible that the audio is popular because of the podcaster level or the audio content characteristics. Therefore, we further validate the effects on audio characteristics and podcaster level. Table 3.9 examines the effects based on the condition of audio characteristics NO.. The results reveal the effect is more pronounced in audios with fewer characteristics (β=0.655 and β =-0.297) (see Table 3.9 columns 1 and 3), which also supports H3 and H4. Table 3.10 compares the results of high-level podcaster and low-level podcaster conditions. The podcaster level is classified based on the podcaster level median. The results reveal that the effect is more pronounced in low-level podcaster audios, supporting H3 and H4. In low-level podcaster situations, Table 3.10 reveals users who are first exposed to a human voice increase the average user evaluation by 32.6%, while users who are first exposed to an AI voice decrease the average user evaluation by 41.3%. In addition, we classify the sample based on whether the other podcaster level is higher than the podcaster level currently listening to. It finds that if the other podcaster level is higher than the podcaster level currently listening to, the effect is pronounced. For users who are first exposed to a human voice, the average user evaluation increases by 54.6% (Table 3.11 column 1). For users who are first exposed to a human voice, the average user evaluation decreases by 41.3% (Table 3.11 column 3). These results (Table 3.11), in addition to validating the last results (Table 3.10), also validate H1 and H2. In a highly achievement-oriented society like ours, the social comparison theory (Festinger, 1954) points out that a person has a primitive drive to compare his/her skill to that of someone he feels to be of slight greater ability. Later, the social comparison theory applied to luxury goods purchases (Pillai and Nair, 2021), received service (Yang and Oliver, 2010), and so on. When a listener finds that the level of another podcaster is higher than the level of the podcaster, he/she is listening to, this triggers a comparison and thus reinforces the comparison effect. It shows that the contrast effect is significant when individuals are comparing, which validates the underlying mechanism.

Table 3.8. Subsample Analysis: Conditional on the Audio Play Count

Dependent Variable:	Review Positivity			
Hybrid Status:	Human->Hybrid		AI->Hybrid	
PlayCount:	Fewer	More	Fewer	More
Hybrid Service*After	**0.504****	**-0.134**	**-0.217***	**0.158**
	(2.388)	**(-0.783)**	**(-1.952)**	**(0.481)**
After	-0.029	0.260*	0.239**	-0.302
	(-0.214)	(1.691)	(2.278)	(-0.915)
Hybrid Service	-0.496***	-0.058	-0.278**	-0.169
	(-3.667)	(-0.579)	(-2.007)	(-1.576)
Ln (NLP Accuracy)	0.011	-0.204**	-0.223***	-0.137***
	(0.382)	(-2.427)	(-4.893)	(-3.695)
Ln (Review No.)	0.043	0.077**	-0.008	-0.027
	(0.942)	(2.140)	(-0.150)	(-0.450)
Single Podcaster	0.053	-	-0.056	-0.070
	(0.555)	-	(-0.811)	(-0.705)
Topics	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Constant	0.686***	0.281**	0.442***	0.552***
	(7.580)	(2.290)	(3.002)	(3.421)
N	187	155	272	192
Adjusted r2	0.118	0.087	0.136	0.119

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.9. Subsample Analysis: Conditional on the Audio Characteristics NO.

Dependent Variable:	Review Positivity			
Hybrid Status:	Human->Hybrid		AI->Hybrid	
Characteristics NO.	Fewer	More	Fewer	More
Hybrid Service*After	**0.655**********	**0.060**	**-0.297***	**-0.105**
	(3.198)	**(0.341)**	**(-1.690)**	**(-0.795)**
After	-0.205*	0.188	0.213	0.112
	(-1.800)	(1.141)	(1.228)	(0.892)
Hybrid Service	-0.723**	-0.621***	-0.420**	-0.288***
	(-2.264)	(-3.390)	(-2.010)	(-3.116)
Ln (NLP Accuracy)	-0.018	-0.031	-0.183***	-0.193***
	(-0.486)	(-0.885)	(-3.372)	(-3.955)
Ln (Review No.)	-0.013	0.057*	-0.002	0.017
	(-0.174)	(1.793)	(-0.042)	(0.288)
Single Podcaster	0.404	0.151	0.065	-0.082
	(1.621)	(0.734)	(0.809)	(-1.090)
Topics	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Constant	1.167***	0.890***	0.598***	0.555***
	(3.788)	(4.603)	(2.718)	(4.131)
N	103	239	170	294
Adjusted r2	0.223	0.216	0.104	0.099

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.10. Subsample Analysis: Conditional on the Podcaster level

Dependent Variable:	Review Positivity			
Hybrid Status:	Human->Hybrid		AI->Hybrid	
Podcaster Level	Low	High	Low	High
Hybrid Service*After	**0.326***	**0.124**	**-0.413*****	**0.303**
	(1.702)	**(0.456)**	**(-3.726)**	**(1.150)**
After	-0.131	0.100	0.236**	-0.081
	(-1.138)	(0.384)	(2.188)	(-0.300)
Hybrid Service	-0.278	-0.611***	0.0722	-0.289***
	(-1.156)	(-8.157)	(0.926)	(-3.935)
Ln (NLP Accuracy)	-0.007	-0.022	-0.211***	-0.179***
	(-0.197)	(-0.445)	(-4.355)	(-3.218)
Ln (Review No.)	0.025	0.037	-0.066	0.035
	(0.544)	(0.613)	(-1.230)	(0.550)
Single Podcaster	-	-	-	0.071
	-	-	-	(0.727)
Genre	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Constant	0.692*	0.921***	0.235***	0.418***
	(1.798)	(9.252)	(2.707)	(3.846)
N	163	179	248	216
Adjusted r2	0.127	0.215	0.149	0.151

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.11. Subsample Analysis: Conditional on the other Podcaster Level

Dependent Variable:	Review Positivity			
Hybrid Status:	Human->Hybrid		AI->Hybrid	
Is the other podcaster better?	Yes	No	Yes	No
Hybrid Service*After	**0.546***	**0.123**	**-0.413*****	**0.303**
	(1.870)	**(0.652)**	**(-3.726)**	**(1.150)**
After	-0.017	-0.009	0.236**	-0.081
	(-0.067)	(-0.057)	(2.188)	(-0.300)
Hybrid Service	-0.318	-0.417***	0.072	-0.289***
	(-1.102)	(-2.841)	(0.926)	(-3.935)
Ln (NLP Accuracy)	-0.078**	0.008	-0.211***	-0.179***
	(-2.207)	(0.238)	(-4.355)	(-3.218)
Ln (Review No.)	-0.002	0.059**	-0.066	0.035
	(-0.022)	(1.976)	(-1.230)	(0.550)
Single Podcaster	-0.174	0.128	-	-
	(-0.657)	(0.685)	-	-
Topics	Yes	Yes	Yes	Yes
Characteristics	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes
Constant	0.718***	0.604***	0.235***	0.418***
	(4.481)	(2.860)	(2.707)	(3.846)
N	79	263	248	216
Adjusted r2	0.166	0.171	0.149	0.151

t statistics with robust standard errors in parentheses;
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.8. Conclusion and Discussions

In this research, we report a robust association between the introduction of a human-AI hybrid service and the average user evaluation of the service in the context of an online streaming digital platform. Further, we reveal that the effect made by the introduction of a human-AI hybrid service on user evaluations also depends on the service content's popularity. We found the effect is pronounced if the audio is less popular. An audio with high-level popularity will divert listeners' attention to contrast human and AI services. AI technology has become an essential tool for digital platforms. However, the impact of AI on user evaluations of the service is not fully understood. This online streaming platform introduced the AI podcaster service at the end of 2020, which provides a great opportunity for us to investigate the effects of human-AI hybrid service. Especially, we divide the process of human-AI hybrid service introduction into two situations from a temporal perspective. The platform offers human services first (Human service → Human-AI services), and the platform offers AI services first (AI service → Human-AI services). We reveal that the contrast effect plays a vital role in the process of the human-AI hybrid service introduction. AI voice listeners' average user evaluation of the service is negatively associated with the later introduction of a human-AI hybrid service (decrease 19.8%), and human voice listeners' user evaluations of the service is positively associated with the later introduction of a human-AI hybrid service (increase 34.8%). We explain it using the contrast effect, which will enhance the difference between the human voice and AI voice services. In addition, the contrast effect is more pronounced if the audio is less popular. We also confirm users generally consider human service to be better (Luo et al., 2019, Mou and Xu, 2017).

This study makes contributions to the Information System literature in three aspects. First, this study helps us to understand a new phenomenon of human-AI coexistence on digital platforms from a temporal perspective. Second, in academia, most literature studied this topic in two streams—system design (e.g., Correia et al., 2020; Ostheimer et al., 2021) and compare the human-only mode, the AI-only mode, and the human-AI coexistence mode (e.g., Peeters et al., 2021; Sowa et al., 2021). The user evaluation of the service of the human-AI hybrid service is still not thoroughly understood. This research incorporates the human-AI hybrid service introduction process and reveals that the introduction of other types of service influences the user evaluation of the service. It sheds light on the underlying mechanism of the effect of human-AI hybrid service on the average user evaluation of the service. We find that the contrast effect is dominant in the situation, compared to the first mover advantage. With the introduction of different types of services, the average user evaluation of AI services becomes worse, while the average user evaluation of human services becomes better. Third, this research contributes to AI on digital platforms literature and AI introduction in business literature. Existing AI on digital platform literature mainly studied how AI supports digital platform processes (Alt, 2021; Mucha and Seppälä, 2020) and AI as a service on digital platforms (Alt, 2021; Colace et al., 2018). In AI in business research, most research studied the antecedents (i.e., Baabdullah et al., 2021; Holmström, 2022; Jöhnk et al., 2021; Ostrom et al., 2019) and consequences (i.e., Luo et, al., 2019; Schuetzler et al., 2020; Shi et al., 2020) of introduction AI applications in business. Little research studied human and AI services as a whole, although the human-AI hybrid service is gaining momentum.

In practice, our research provides a first step toward understanding the effects of the human-AI hybrid service introduction process in digital platforms. This research helps digital platform managers by reminding managers that the service implementation strategies affect user evaluations of the service. If the platform is offering an AI service, the introduction of a human service will be negatively associated with AI service user evaluations. In such cases, platform managers need to consider the user evaluation of the service more carefully. There may be countermeasures that can be taken to avoid degradation of the average user evaluation, such as upgrading the AI service at the same time. In addition, the user evaluation of the service has long been a concern for digital platform managers. This study takes a user evaluation perspective, which can help the development of digital platforms and give platform managers a better understanding of what their users think. Good user evaluations of the service ultimately help the digital platform to attract more users and helps the digital platform stand out from the competitors.

This research subjects to several limitations. First, the study was conducted in the context of the online streaming platform and, therefore, the findings may not be applicable to other industries. In AI first situation, the podcaster level and hybrid service have a high correlation, which makes it hard for us to control podcaster level in models. Due to the GPU resources limitations, we cannot train Large Language Models. Other features of review can be captured in the future.

Chapter 4 : Conclusions and Future Research

AI has been demonstrated can carry out very complex tasks and help people in various aspects. It is viewed as critical to society's digital transformation, and it has become a top priority for the EU and China. However, users' human-AI interaction experience may push the AI applications back and damage the enormous value of AI. Therefore, in addition to paying attention to the technological development of AI, the research on human-AI interaction also needs attention. In order to enrich the understanding of human-AI interaction, the two studies investigated human-AI interaction from different perspectives. The first study investigated the causes and effects of the human-AI conversational process—conversation inefficiency. The second study focuses on the effects of introducing human-AI hybrid services to digital platforms.

To advance the understanding of human-AI interaction, scholars from different disciplines such as Communication, Education, and Organization joined AI research and highlighted AI research is no longer computer science research. However, the research in human-AI interaction is still considerably limited, needing further effort to expand, especially in the Information Systems field. Accordingly, the two research in this thesis study how chatbots' social designs affect the human-AI conversational process and, therefore, influence associated outcomes (study 1) and how the average user evaluation of the service changes in the process of introducing a human-AI hybrid service (study 2). In the first research, we found chatbot proactivity design is essential in reducing cognitive uncertainty and helping human-AI communication inefficiency. As a result, reduced communication inefficiency increases customers' satisfaction and task success. The social identity design, which adopted a relational communication strategy, was not found a direct effect on communication inefficiency but play as a moderator to weaken the negative effect of communication inefficiency on customer satisfaction. The second research studied customers' user evaluations of the services about the human-AI hybrid service from the perspective of the contrast effect. This research found that the contrast effect is dominant in the process of introducing a human-AI hybrid service. By comparing AI voice service and human voice service, the gap between the two services will be highlighted, thus polarizing the average user evaluation of the service. AI service user evaluations became worse, while human service user evaluations became better.

The two research provide foundations for future research. First, we extend human-AI interaction research into the human-AI conversational process. In the future, more conversational processes will be studied, such as how human-AI conversation topics change and how human-AI conversation conflicts happen. In addition, not just limited to human-AI conversational processes, future research could extend to physical interaction processes. Second, it helps human-AI scholars take a first step toward understanding the effects of human-AI hybrid services from the users' perspective. Future research could be extended to other cases of human-AI hybrid service situations, such as human-centered AI situations. Another type of human-AI hybrid is also deserved to be studied, such as human-AI hybrid intelligence. The effects made by the human-AI hybrid also can be examined in more contexts.

References

Adam, M., Wessel, M., and Benlian, A. (2020). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*. Published. https://doi.org/10.1007/s12525-020-00414-7.

Adams, S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., Hall, J. S., Samsonovich, A., Scheutz, M., Schlesinger, M., Shapiro, S. C., and Sowa, J. (2012). Mapping the Landscape of Human-Level Artificial General Intelligence. *AI Magazine*, 33(1), 25-42. https://doi.org/10.1609/aimag.v33i1.2322.

Agarwal, R., and Prasad, J. (1998). A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. *Information Systems Research*, *9*(2), 204–215. https://doi.org/10.1287/isre.9.2.204.

Agomuoh, F. (2023, January 27). *The 6 biggest problems with CHATGPT right now*. Digital Trends. https://www.digitaltrends.com/computing/the-6-biggest-problems-with-chatgpt-right-now/.

Akhtar, M., Neidhardt, J., and Werthner, H. (2019). The potential of chatbots: analysis of chatbot conversations. In *2019 IEEE 21st Conference on Business Informatics (CBI)* (Vol. 1, pp. 397-404). IEEE.

Alekseeva, L., Gine, M., Samila, S., & Taska, B. (2021). AI adoption and firm performance: Management versus it. *Academy of Management Proceedings*, *2021*(1), 15935. https://doi.org/10.5465/ambpp.2021.15935abstract.

Alsheibani, S., Cheung, Y., & Messom, C. (2018). Artificial Intelligence Adoption: AI-readiness at Firm-Level. *PACIS*, 4, 231-245.

Alt, R. (2021). Electronic markets on digital platforms and ai. *Electronic Markets*, *31*(2), 233–241. https://doi.org/10.1007/s12525-021-00489-w.

Altman, I., and Taylor, D.A. (1973) Social Penetration: The Development of Interpersonal Relationships. *Holt, Rinehart, and Winston*, New York, 459.

Andalib, T. W., Darun, M. R., & Azizan, A. (2013). Frustration of employees: Reasons, dimensions and resolving techniques. *WCIK E Journal of Integration Knowledge*, 15-25.

Antheunis, M. L., Schouten, A. P., Valkenburg, P. M., and Peter, J. (2011). Interactive Uncertainty Reduction Strategies and Verbal Affection in Computer-Mediated Communication. *Communication Research*, 39(6), 757–780. https://doi.org/10.1177/0093650211410420.

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior, 85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051.

Ashktorab, Z., Jain, M., Liao, Q. V., and Weisz, J. D. (2019). Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3290605.3300484.

Baabdullah, A. M., Alalwan, A. A., Slade, E. L., Raman, R., & Khatatneh, K. F. (2021). SMEs and artificial intelligence (AI): Antecedents and consequences of AI-based B2B practices. Industrial Marketing Management, 98, 255-270.

Bahtiri, L., & Peeters, D. G. T. (2019). The Role of Uncertainty and Humor in Friendship Formation with Social Chatbots, Tilburg University.

Balcerzak, B., Jaworski, W., & Wierzbicki, A. (2014, August). Application of TextRank algorithm for credibility assessment. *In 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)* (Vol. 1, pp. 451-454). IEEE.

Bansal, G., Nushi, B., Kamar, E., Lasecki, W.S., Weld, D.S., and Horvitz, E. (2019a). Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. *AAAI 2019*.

Bansal, G., Nushi, B., Kamar, E., Weld, D. S., Lasecki, W. S., and Horvitz, E. (2019b). Updates in Human-AI Teams: Understanding and Addressing the Performance/Compatibility Tradeoff. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 2429-2437. https://doi.org/10.1609/aaai.v33i01.33012429.

Bauman, K., Liu, B., & Tuzhilin, A. (2016, September). Recommending Items with Conditions Enhancing User evaluation on the services Based on Sentiment Analysis of Reviews. *In CBRecSys@ RecSys* (pp. 19-22).

Beattie, A., Edwards, A. P., and Edwards, C. (2020). A Bot and a Smile: Interpersonal Impressions of Chatbots and Humans Using Emoji in Computer-mediated Communication. *Communication Studies*, *71*(3), 409–427. https://doi.org/10.1080/10510974.2020.1725082.

Berger, C. R., & Bradac, J. J. (1982). Language and social knowledge: Uncertainty in interpersonal relations (Vol. 2). Hodder Education.

Berger, C. R., and Calabrese, R. J. (1975). Some Explorations In Initial Interaction And Beyond: Toward A Developmental Theory Of Interpersonal Communication. Human Communication Research, 1(2), 99-112. doi:10.1111/j.1468-2958.1975.tb00258.x

Blau, Peter (1964).*Exchange and power in social life*. New York: Wiley.ISBN 978-0-88738-628-2.

Brown, C., Story, G. W., Mourão-Miranda, J., and Baker, J. T. (2021). Will artificial intelligence eventually replace psychiatrists? In *British Journal of Psychiatry* (Vol. 218, Issue 3, pp. 131–134). Cambridge University Press. https://doi.org/10.1192/bjp.2019.245.

Brynjolfsson, E., Hui, X., and Liu, M. (2019). Does machine translation affect international trade? Evidence from a large digital platform. *Management Science*, *65*(12), 5449-5460.

Burgener, A. M. (2020). Enhancing Communication to Improve Patient Safety and to Increase Patient Satisfaction. *The Health Care Mana*ger, 39(3), 128–132. https://doi.org/10.1097/hcm.0000000000000298.

Calo, R. (2017). Artificial Intelligence Policy: A roadmap. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3015350.

Cameron, J., Hurd, G. A., Schwarzenegger, A., Biehn, M., Hamilton, L., Winfield, P., Henriksen, L., ... MGM Home Entertainment Inc. (2004).*The Terminator*. Santa Monica, CA: MGM DVD.

Candello, H., Pinhanez, C., and Figueiredo, F. (2017). Typefaces and the Perception of Humanness in Natural Language Chatbots. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3025453.3025919

Candello, H., Pinhanez, C., Pichiliani, M., Cavalin, P., Figueiredo, F., Vasconcelos, M., and Do Carmo, H. (2019). The effect of audiences on the user evaluation on the service with conversational interfaces in physical spaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Cem Delmegani. (2023). *9 epic chatbot/conversational bot failures (2023 update)*. AIMultiple. https://research.aimultiple.com/chatbot-fail/#bots-that-dont-accept-no-for-an-answer

Chaves, A. P., and Gerosa, M. A. (2020). How Should My Chatbot Interact? A Survey on Social Characteristics in Human–Chatbot Interaction Design. *International Journal of Human–Computer Interaction*, 37(8), 729–758. https://doi.org/10.1080/10447318.2020.1841438.

Chen, H., Chan-Olmsted, S., Kim, J., & Sanabria, I. M. (2021). Consumers' perception on artificial intelligence applications in marketing communication. Qualitative Market Research: An International Journal, 25(1), 125-142.

Cho, E. (2019). Hey google, can I ask you something in private? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3290605.3300488.

Cho, E., Molina, M. D., and Wang, J. (2019). The Effects of Modality, Device, and Task Differences on Perceived Human Likeness of Voice-Activated Virtual Assistants. *Cyberpsychology, Behavior, and Social Networking*, 22(8), 515–520. https://doi.org/10.1089/cyber.2018.0571.

Chung, M., Ko, E., Joung, H., and Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*, *117*, 587–595. https://doi.org/10.1016/j.jbusres.2018.10.004.

Colace, F., Santo, M. D., Lombardi, M., Pascale, F., Pietrosanto, A., and Lemma, S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, 528–533. https://doi.org/10.18178/ijmerr.7.5.528-533.

Correia, A., Jameel, S., Schneider, D., Paredes, H., and Fonseca, B. (2020). A Workflow-Based Methodological Framework for Hybrid Human-AI Enabled Scientometrics. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, 2876–2883. https://doi.org/10.1109/BigData50022.2020.9378096

Corti, K., and Gillespie, A. (2016). Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human. *Computers in Human Behavior*, *58*, 431–442. https://doi.org/10.1016/j.chb.2015.12.039.

David Fry (2016). "Five Reasons for Proactive Communications". *Inside Indiana Business*. Retrieved July 6, 2016.

De Dreu, C. K. W., and Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. Journal of Applied Psychology, 88(4), 741–749. https://doi.org/10.1037/0021-9010.88.4.741

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., and Ebel, P. (2019). The future of human-ai collaboration: A taxonomy of design knowledge for hybrid intelligence systems. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.24251/hicss.2019.034.

Dellermann, D., Calma, A., Lipusch, N., Weber, T., Weigel, S., and Ebel, P. (2021). *The Future of Human-AI Collaboration: A Taxonomy of Design Knowledge for Hybrid Intelligence Systems*.

Deng, X., Guo, X., Wu, Y. J., and Chen, M. (2021). Perceived Environmental Dynamism Promotes Entrepreneurial Team Member's Innovation: Explanations Based on the Uncertainty Reduction Theory. *International Journal of Environmental Research and Public Health*, *18*(4), 2033. https://doi.org/10.3390/ijerph18042033.

Du Boulay, B. (2016). Artificial intelligence as an effective classroom assistant. IEEE *Intelligent Systems*, 31(6), 76-81.

Dubey, A., Abhinav, K., Jain, S., Arora, V., and Puttaveerana, A. (2020, February 27). HACO: A Framework for Developing Human-AI Teaming. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3385032.3385044

Edstrom, A. (1977). User Influence and the Success of MIS Projects: A Contingency Approach. *Human Relations*, *30*(7), 589–607. https://doi.org/10.1177/001872677703000702.

Edwards, A., Edwards, C., Spence, P. R., Harris, C., and Gambino, A. (2016). Robots in the classroom: Differences in students' perceptions of credibility and learning between “teacher as robot” and “robot as teacher.” *Computers in Human Behavior*, *65*, 627–634. https://doi.org/10.1016/j.chb.2016.06.005

Eeuwen, M. van. (2017, February 8). Mobile conversational commerce: messenger chatbots as the next interface between businesses and consumers [Info:eu-repo/ semantics/masterThesis]. http://essay.utwente.nl/71706/.

Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2), 117-140.

Fournier-Tombs, E. (2021). Towards a United Nations Internal Regulation for Artificial Intelligence. *Big Data and Society*, *8*(2). https://doi.org/10.1177/20539517211039493.

Frick, N. R., Mirbabaie, M., Stieglitz, S., & Salomon, J. (2021). Maneuvering through the stormy seas of digital transformation: the impact of empowering leadership on the AI readiness of enterprises. *Journal of Decision Systems*, 30(2-3), 235-258.

Gao, Z., and Jiang, J. (2021). Evaluating Human-AI Hybrid Conversational Systems with Chatbot Message Suggestions. *International Conference on Information and Knowledge Management, Proceedings*, 534–544. https://doi.org/10.1145/3459637.3482340

Giannelis, M. (2023, March 15). *Why customers don't like or hate chatbots - annoying & bad for business* ... Retrieved April 11, 2023, from https://www.techbusinessnews.com.au/blog/why-customers-dont-like-or-hate-chatbots-bad-for-business/

Gill, T. (2020). Blame It on the Self-Driving Car: How Autonomous Vehicles Can Alter Consumer Morality. *Journal of Consumer Research*, 47(2), 272-291. doi:10.1093/jcr/ucaa018

Gnewuch, U., Morana, S., Adam, M.T., and Maedche, A. (2018). Faster is Not Always Better: Understanding the Effect of Dynamic Response Delays in Human-Chatbot Interaction. *ECIS*.

Grant, A. M., & Ashford, S. J. (2008). The dynamics of proactivity at work. Research in organizational behavior, 28, 3-34.

Hecht, M. L. (1978). Measures of communication satisfaction. *Human Communication Research*, 4(4), 350–368. https://doi.org/10.1111/j.1468-2958.1978.tb00721.x.

Herr, P. (1986). Consequences of priming: judgment and behavior. J. Pers. Soc. Psychol. 51, 1106–1115. doi: 10.1037/0022-3514.51.6.1106.

Hill, J., Randolph Ford, W., and Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior*, *49*, 245–250. https://doi.org/10.1016/j.chb.2015.02.026.

Hofmann, P., Samp, C., and Urbach, N. (2020). Robotic process automation. *Electronic Markets*, 30(1), 99–106. https://doi.org/10.1007/ s12525-019-00365-8.

Holmström, J. (2022). From AI to digital transformation: The AI readiness framework. *Business Horizons*, 65(3), 329-339.

Holstein Kenneth and Aleven, V. and R. N. (2020). A Conceptual Framework for Human–AI Hybrid Adaptivity in Education. In M. and M. K. and L. R. and M. E. Bittencourt Ig Ibert and Cukurova (Ed.), *Artificial Intelligence in Education* (pp. 240–254). Springer International Publishing.

Hovland, C. I., Harvey, O. J., and Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *The Journal of Abnormal and Social Psychology, 55*(2), 244–252. https://doi.org/10.1037/h0048480.

Huang, C.-Z. A., Koops, H. V., Newton-Rex, E., Dinculescu, M., and Cai, C. J. (2020). *AI Song Contest: Human-AI Co-Creation in Songwriting*. http://arxiv.org/abs/2010.05388

Huang, M. H., and Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459.

Humphreys, M. S., and Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review*, 91(2), 153–184. doi: 10.1037/0033-295x.91.2.153.

Hussain, Z. I., Sivarajah, U., and Hussain, N. (2019). The role of a digital engineering platform in appropriating the creation of new work-related mind-set and organisational discourse in a large multi-national company. *International Journal of Information Management*, *48*, 218–225. https://doi.org/10.1016/j.ijinfomgt.2018.09.014.

Illescas-Manzano, M. D., Vicente López, N., Afonso González, N., and Cristofol Rodríguez, C. (2021). Implementation of chatbot in online commerce, and open innovation. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(2), 125.

Intellexer. (2019, January 22). *Customer Support Agent vs Chatbot. Pros and Cons of Both of Them*. Retrieved from https://chatbotslife.com/customer-support-agent-vs-chatbot-pros-and-cons-of-both-of-them-ca92957b89d2.

Jacobs, M. A., Yu, W., and Chavez, R. (2016). The effect of internal communication and employee satisfaction on supply chain integration. *International Journal of Production Economics*, 171, 60–70. https://doi.org/10.1016/j.ijpe.2015.10.015.

Jain, M., Kumar, P., Kota, R., and Patel, S. N. (2018). Evaluating and Informing the Design of Chatbots. *Proceedings of the 2018 Designing Interactive Systems Conference*. Published. https://doi.org/10.1145/3196709.3196735.

Jarek, K., and Mazurek, G. (2019). Marketing and artificial intelligence. *Central European Business Review*, *8*(2), 46–55. https://doi.org/10.18267/j.cebr.213.

Jenkins, M., Churchill, R., Cox, S., and Smith, D. (2007). Analysis of User Interaction with Service Oriented Chatbot Systems. Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments Lecture Notes in Computer Science,76-83. doi:10.1007/978-3-540-73110-8_9.

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., and Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. In *Stroke and Vascular Neurology* (Vol. 2, Issue 4, pp. 230–243). BMJ Publishing Group. https://doi.org/10.1136/svn-2017-000101.

Jin, C., Chen, W., Cao, Y., Xu, Z., Tan, Z., Zhang, X., Deng, L., Zheng, C., Zhou, J., Shi, H., and Feng, J. (2020). Development and evaluation of an artificial intelligence system for COVID-19 diagnosis. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-18685-1.

Jin, Y., Cai, W., Chen, L., Htun, N. N., and Verbert, K. (2019). MusicBot: Evaluating Critiquing-Based Music Recommenders with Conversational Interaction. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. https://doi.org/10.1145/3357384.3357923.

Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or not, AI comes—an interview study of organizational AI readiness factors. Business & Information Systems Engineering, 63, 5-20.

Johnson, A.; Proctor, R. (2003). Attention (1st ed.). SAGE Publications. Retrieved from https://www.perlego.com/book/1004620/attention-theory-and-practice-pdf (Original work published 2003)

Johnson, D. G., & Verdicchio, M. (2017). Ai anxiety. *Journal of the Association for Information Science and Technology*, *68*(9), 2267–2270. https://doi.org/10.1002/asi.23867.

Jonze, S., Ellison, M., Landay, V., Lupi, D., Farrey, N., Barnard, C., Phoenix, J., ... Warner Bros. Entertainment. (2014). Her.

Kahn, L. H., Savas, O., Morrison, A., Shaffer, K. A., and Zapata, L. (2020). Modelling Hybrid Human-Artificial Intelligence Cooperation: A Call Center Customer Service Case Study. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, 3072–3075. https://doi.org/10.1109/BigData50022.2020.9377747

Karaca, O., Çalışkan, S. A., & Demir, K. (2021). Medical artificial intelligence readiness scale for medical students (MAIRS-MS)–development, validity and reliability study. BMC medical education, 21, 1-9.

Katzenstein, P. (1996). *The Culture of National Security (Illustrated ed.)*. Columbia University Press.

Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T., and Bernstein, M. S. (2020). Conceptual Metaphors Impact Perceptions of Human-AI Collaboration. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2). https://doi.org/10.1145/3415234

Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2020). My teacher is a machine: Understanding students' perceptions of AI teaching assistants in online education. International Journal of Human–Computer Interaction, 36(20), 1902-1911.

Kim, S., Chen, R. P., and Zhang, K. (2016). Anthropomorphized Helpers Undermine Autonomy and Enjoyment in Computer Games. *Journal of Consumer Research*, *43*(2), 282–302. https://doi.org/10.1093/jcr/ucw016.

Komatsu, T., Kobayashi, K., Yamada, S., Funakoshi, K., and Nakano, M. (2012). A Comparison of Artificial Subtle Expressions with Human-like Expressions on Expressing Accuracy Level. *Transactions of the Japanese Society for Artificial Intelligence*, *27*(5), 263–270. https://doi.org/10.1527/tjsai.27.263.

Korteling, J. E. (Hans), van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., and Eikelboom, A. R. (2021). Human- versus Artificial Intelligence. *Frontiers in Artificial Intelligence*, *4*. https://doi.org/10.3389/frai.2021.622364

Kull, A. J., Romero, M., and Monahan, L. (2021). How may I help you? Driving brand engagement through the warmth of an initial chatbot message. *Journal of Business Research*, *135*, 840–850. https://doi.org/10.1016/j.jbusres.2021.03.005.

Kushner, Laura H. (2008). *Contrast in judgments of mental health*. p. 1. ISBN 978-0-549-91314-6. Retrieved 24 March 2011.

Lai, V., Carton, S., Bhatnagar, R., Liao, Q. V., Zhang, Y., and Tan, C. (2022). Human-ai collaboration via Conditional Delegation: A case study of content moderation. *CHI Conference on Human Factors in Computing Systems.* https://doi.org/10.1145/3491102.3501999.

Lai, Y., Kankanhalli, A., and Ong, D. (2021). Human-AI Collaboration in Healthcare: A Review and Research Agenda. *HICSS*.

Leah. (2018, November 20). *5 Steps to a Catchy Bot Name*. Retrieved September 02, 2020, from https://www.userlike.com/en/blog/bot-names.

Lee, K. M., Peng, W., Jin, S. A., and Yan, C. (2006). Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human–Robot Interaction. *Journal of Communication*, *56*(4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x.

Lee, K.-F. (2021). A human blueprint for AI coexistence. In J. Von Braun, M. Archer, G. Reichberg, and M. S. Sorondo (Eds.), Robotics, AI, and humanity: Science, ethics, and policy (pp. 261–269). Springer.

Li, C. H., Yeh, S. F., Chang, T. J., Tsai, M. H., Chen, K., and Chang, Y. J. (2020). A Conversation Analysis of Non-Progress and Coping Strategies with a Banking Task-Oriented Chatbot. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Published. https://doi.org/10.1145/3313831.3376209.

Li, W., Zhou, X., and Yang, Q. (2021). Designing medical artificial intelligence for in- and out-groups. *Computers in Human Behavior*, *124*, 106929. https://doi.org/10.1016/j.chb.2021.106929.

Li, X. (2021). The new development direction of artificial intelligence - Human cyber physical intelligence. *2021 10th International Conference on Educational and Information Technology, ICEIT 2021*, 249–252. https://doi.org/10.1109/ICEIT51700.2021.9375623

Liu, B. (2021). In ai we trust? effects of agency locus and transparency on uncertainty reduction in human–ai interaction. *Journal of Computer-Mediated Communication*, *26*(6), 384–402. https://doi.org/10.1093/jcmc/zmab013

Logacheva, V., Burtsev, M., Malykh, V., Polulyakh, V., and Seliverstov, A. (2018). ConvAI Dataset of Topic-Oriented Human-to-Chatbot Dialogues. *The NIPS '17 Competition: Building Intelligent Systems*, 47–57. https://doi.org/10.1007/978-3-319-94042-7_3.

Longoni, C., Bonezzi, A., and Morewedge, C. K. (2019). Resistance to Medical Artificial Intelligence. *Journal of Consumer Research*, *46*(4), 629–650. https://doi.org/10.1093/jcr/ucz013.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *The Journal of finance*, 50(1), 23-51.

Lu LX. (2017). *The Power of Customization: Giving the Customer Exactly What They Want*. Retrieved from https://www.bluleadz.com/blog/the-power-of-customization-giving-the-customer-exactly-what-they-want.

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. Computers and Education: Artificial Intelligence, 3, 100076.

Luo, X., Tong, S., Fang, Z., and Qu, Z. (2019). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*. Published. https://doi.org/10.1287/mksc.2019.1192.

Lurings, N. J. (2019). The role of intimacy of self-disclosure and perceived similarity in the development of social attraction in human-human CMC and human-chatbot interactions, Tilburg University.

Malgonde, O., Zhang, H., Padmanabhan, B., and Limayem, M. (2020). Taming the complexity in search matching: Two-sided recommender systems on digital platforms. *MIS Quarterly*, *44*(1), 48–84. https://doi.org/10.25300/misq/2020/14424.

Manis, M., Nelson, T., and Shedler, J. (1988). Stereotypes and social judgment: extremity, assimilation, and contrast. J. Person. Soc. Psychol. 55, 28–36. doi: 10.1037/0022-3514.55.1.28.

McTear, M., Callejas, Z., and Griol, D. (2016). Creating a conversational interface using chatbot technology. In The conversational interface (pp. 125-159). Springer, Cham.

Mirbabaie, M., Brünker, F., Möllmann, N. R., & Stieglitz, S. (2022). The rise of artificial intelligence–understanding the AI identity threat at the workplace. *Electronic Markets*, 1-27.

Miroshnichenko, A. (2018). AI to Bypass Creativity. Will Robots Replace Journalists? (The Answer Is "Yes"). Inf., 9, 183.

Mostafa, R. B., and Kasamani, T. (2021). Antecedents and consequences of chatbot initial trust. *European Journal of Marketing*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/ejm-02-2020-0084.

Mou, Y., and Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72, 432–440. https://doi.org/10.1016/j.chb.2017.02.067.

Mucha, T., and Seppälä, T. (2020). Artifcial intelligence platforms–a new research agenda for digital platform economy. ETLA Working Papers No 76, http://pub.etla.f/ETLA-Working-Papers-76.pdf, accessed May 11, 2021.

Mueller, S. T., Veinott, E. S., Hoffman, R. R., Klein, G., Alam, L., Mamun, T., and Clancey, W. J. (2020). *Principles of Explanation in Human-AI Systems*. www.aaai.org.

Muldoon, O. T., Haslam, S. A., Haslam, C., Cruwys, T., Kearns, M., & Jetten, J. (2019). The social psychology of responses to trauma: Social identity pathways associated with divergent traumatic responses. European Review of Social Psychology, 30(1), 311-348.

Muthén, L.K. and Muthén, B.O. (1998-2010). Mplus User's Guide. Sixth Edition. Los Angeles, CA: Muthén and Muthén.

Myers, C., Furqan, A., Nebolsky, J., Caro, K., and Zhu, J. (2018). Patterns for How Users Overcome Obstacles in Voice User Interfaces. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3173574.3173580.

Najdawi, A. (2020, July). Assessing AI readiness across organizations: The case of UAE. *In 2020 11th International Conference on Computing, Communication and Networking Technologies* (ICCCNT) (pp. 1-5). IEEE.

Nelson, G. L. (1992). The relationship between the use of personal, cultural examples in international teaching assistants' lectures and uncertainty reduction, student attitude, student recall, and ethnocentrism. *International Journal of Intercultural Relations*, *16*(1), 33–52. https://doi.org/10.1016/0147-1767(92)90004-e.

Neuliep, J. W. (2012). The Relationship among Intercultural Communication Apprehension, Ethnocentrism, Uncertainty Reduction, and Communication Satisfaction during Initial Intercultural Interaction: An Extension of Anxiety and Uncertainty Management (AUM) Theory. *Journal of Intercultural Communication Research*, 41(1), 1–16. https://doi.org/10.1080/17475759.2011.623239.

Nolan, J., Joy, L., Wood, E. R., Newton, T., Hopkins, A., Wright, J., Harris, E., ... Warner Bros. Entertainment Canada (Firm),. (2017). Westworld: Season one.

Ong, K. K. (2017). *Business Application of ChatBots*. Retrieved from https://chatbotsmagazine.com/business-application-of-chatbots-afb952cfdb93.

Ostheimer, J., Chowdhury, S., and Iqbal, S. (2021). An alliance of humans and machines for machine learning: Hybrid intelligent systems and their design principles. *Technology in Society*, *66*. https://doi.org/10.1016/j.techsoc.2021.101647.

Ostrom, A. L., Fotheringham, D., & Bitner, M. J. (2019). Customer acceptance of AI in service encounters: understanding antecedents and consequences. *Handbook of Service Science*, Volume II, 77-103.

Paek, T., and Horvitz, E. (1999). Uncertainty, utility, and misunderstanding: A decision-theoretic perspective on grounding in conversational systems. In *AAAI Fall Symposium on Psychological Models of Communication*, North.

Park, M., Aiken, M., and Salvador, L. (2018). How do Humans Interact with Chatbots?: An Analysis of Transcripts. *International journal of management and information technology,* 3338–3350. https://doi.org/10.24297/ijmit.v14i0.7921.

Peeters, M. M. M., van Diggelen, J., van den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., and Raaijmakers, S. (2021). Hybrid collective intelligence in a human–AI society. *AI and Society*, *36*(1), 217–238. https://doi.org/10.1007/s00146-020-01005-y.

Peng, Z., Kwon, Y., Lu, J., Wu, Z., and Ma, X. (2019). Design and Evaluation of Service Robot's Proactivity in Decision-Making Support Process. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3290605.3300328.

Pereira, F. D., Pires, F., Fonseca, S. C., Oliveira, E. H., Carvalho, L. S., Oliveira, D. B., and Cristea, A. I. (2021). Towards a human-ai hybrid system for categorising programming problems. *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*. https://doi.org/10.1145/3408877.3432422.

Pierce, R. (2019). AI in Healthcare: Solutions, challenges, and dilemmas in medical decision-making. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3806767.

Pillai, K. G., & Nair, S. R. (2021). The effect of social comparison orientation on luxury purchase intentions. *Journal of Business Research, 134,* 89–100. https://doi.org/10.1016/j.jbusres.2021.05.033.

Pol, H. E. H., Hijman, R., Baare, W. F. C., and van Ree, J. M. (1998). Effects of context on judgments of odor intensities in humans. *Chem. Senses 23*, 131–135. doi: 10.1093/chemse/23.2.131.

Poole, D., Mackworth, A., and Goebel, R. (1998). *Computational Intelligence: A Logical Approach* (1st ed.). Oxford University Press.

Pratt, L., Wiseman, R. L., Cody, M. J., and Wendt, P. F. (1999). Interrogative strategies and information exchange in computer-mediated communication. *Communication Quarterly*, *47*(1), 46–66. https://doi.org/10.1080/01463379909370123.

Prentice, C., Dominique Lopes, S., and Wang, X. (2020). Emotional intelligence or artificial intelligence– an employee perspective. *Journal of Hospitality Marketing and Management*, *29*(4), 377–403. https://doi.org/10.1080/19368623.2019.1647124.

Proost, K., Derous, E., Schreurs, B., Hagtvet, K. A., and De Witte, K. (2008). Selection test anxiety: Investigating applicants' self- vs other-referenced anxiety in a real selection setting. *International Journal of Selection and Assessment*, *16,* 14–26. http://dx.doi.org/10.1111/j.1468-2389.2008 .00405.x.

Ragan, S. L. (1983). A Conversational Analysis of Alignment Talk in Job Interviews. *Annals of the International Communication Association*, *7*(1), 502–516. https://doi.org/10.1080/23808985.1983.11678550.

Rai, A., and Sarker, S. (2019). Editor’s comments Next-Generation Digital Platforms: Toward Human-AI Hybrids. In *MIS Quarterly* (Vol. 43, Issue 1). https://android-developers.googleblog.com/2018/01/how-we-fought-bad-apps-and-malicious.html;.

Raisch, S., and Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, *46*(1), 192–210. https://doi.org/10.5**4**65/AMR.2018.0072.

Raiter, O. (2021). Segmentation of bank consumers for artificial intelligence marketing. *International Journal of Contemporary Financial Issues*, 1(1), 39-54.

Reddy, S., Allan, S., Coghlan, S., and Cooper, P. (2019). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, *27*(3), 491–497. https://doi.org/10.1093/jamia/ocz192.

Redmond, Mark V., "Uncertainty Reduction Theory" (2015). *English Technical Reports and White Papers*. 3. http://lib.dr.iastate.edu/engl_reports/3.

Reicher, S., Spears, R., & Haslam, S. A. (2010). The social identity approach in social psychology. Sage identities handbook, 2010, 45-62.

Robles, J. S. (2017). Misunderstanding as a resource in interaction. Pragmatics. *Quarterly Publication of the International Pragmatics Association (IPrA)*, 27(1), 57–86. https://doi.org/10.1075/prag.27.1.03rob.

Roy, R., and Naidoo, V. (2021). Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *Journal of Business Research*, *126*, 23–34. https://doi.org/10.1016/j.jbusres.2020.12.051.

Saad, B. S. and Fatma C. A. (2016), “Social Interactivity and Its Impact on a User’s Approach Behavior in Commercial Web Sites: A Study Case of Virtual Agent Presence,” *Journal of Marketing Management*, 4 (2), 2333–6099.

Sarris, V. (1967). Adaptation-level theory: two critical experiments on Helson's weighted-average model. Am. J. Psychol. 80, 331–344. doi: 10.2307/1420364.

Schanke, S., Burtch, G., and Ray, G. (2021). Estimating the Impact of "Humanizing" Customer Service Chatbots. *Information Systems Research*, 32(3), 736–751. https://doi.org/10.1287/isre.2021.1015.

Schell, K. L., and Grasha, A. F. (2000). State anxiety, performance accuracy, and work pace in a simulated pharmacy dispensing task. *Perceptual and Motor Skills*, 90, 547–561. http://dx.doi.org/10.2466/pms.2000.90.2.547.

Schlenker, B. R., and Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization model. *Psychological Bulletin*, 92(3), 641-669. doi:10.1037/0033-2909.92.3.641.

Schuetzler, R. M., Grimes, G. M., and Scott Giboney, J. (2020). The impact of chatbot conversational skill on engagement and perceived humanness. *Journal of Management Information Systems*, *37*(3), 875–900. https://doi.org/10.1080/07421222.2020.1790204.

Schuetzler, Ryan and Grimes, Mark and Giboney, Justin and Buckman, Joseph. (2014). Facilitating Natural Conversational Agent Interactions: Lessons from a Deception Experiment.

Schwarzer, R., & Taubert, S. (2002). Tenacious goal pursuits and striving toward personal growth: Proactive coping. *In Beyond coping: Meeting goals, visions and challenges* (pp. 19-35).

Serey, J., Quezada, L., Alfaro, M., Fuertes, G., Vargas, M., Ternero, R., Sabattin, J., Duran, C., and Gutierrez, S. (2021). Artificial intelligence methodologies for data management. In *Symmetry* (Vol. 13, Issue 11). MDPI. https://doi.org/10.3390/sym13112040.

Shi, W., Wang, X., Oh, Y. J., Zhang, J., Sahay, S., and Yu, Z. (2020, April). Effects of Persuasive Dialogues: Testing Bot Identities and Inquiry Strategies. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376843.

Sidelinger, R. J., Nyeste, M. C., Madlock, P. E., Pollak, J., and Wilkinson, J. (2015). Instructor Privacy Management in the Classroom: Exploring Instructors' Ineffective Communication and Student Communication Satisfaction. *Communication Studies*, 66(5), 569–589. https://doi.org/10.1080/10510974.2015.1034875.

Smestad T.L., Volden F. (2019) Chatbot Personalities Matters. In: Bodrunova S. et al. (eds) Internet Science. INSCI 2018. *Lecture Notes in Computer Science*, vol 11551. Springer, Cham. https://doi.org/10.1007/978-3-030-17705-8_15.

Sowa, K., Przegalinska, A., and Ciechanowski, L. (2021). Cobots in knowledge work: Human – AI collaboration in managerial professions. *Journal of Business Research*, *125*, 135–142. https://doi.org/10.1016/j.jbusres.2020.11.038.

Spence, P. R., Westerman, D., Edwards, C., and Edwards, A. (2014). Welcoming our robot overlords: Initial expectations about interaction with a robot. *Communication Research Reports*, *31*(3), 272–280. https://doi.org/10.1080/08824096.2014.924337.

Spielberg, S., Kubrick, S., Kennedy, K., Curtis, B., Watson, I., Osment, H. J., Law, J., ... DreamWorks Home Entertainment (Firm). (2002). A.I. Artificial intelligence.

Spreng, R. A., MacKenzie, S. B., and Olshavsky, R. W. (1996). A Reexamination of the Determinants of Consumer Satisfaction. *Journal of Marketing*, 60(3), 15. https://doi.org/10.2307/1251839.

Stanojević, D., Krstić, M., Jaredić, B., and Dimitrijević, B. (2013). Proactive Coping as a Mediator Between Resources and Outcomes: A Structural Equations Modeling Analysis. *Applied Research in Quality of Life*, *9*(4), 871–885. https://doi.org/10.1007/s11482-013-9274-2.

Stein, J. P., and Ohler, P. (2017). Venturing into the uncanny valley of mind—The influence of mind attribution on the acceptance of humanlike characters in a virtual reality setting. *Cognition*, *160*, 43–50. https://doi.org/10.1016/j.cognition.2016.12.010.

Stella, M., Ferrara, E., and Domenico, M. D. (2018). Bots increase exposure to negative and inflammatory content in online social systems. *Proceedings of the National Academy of Sciences*, 115(49), 12435-12440. doi:10.1073/pnas.1803470115.

Summers, T. A., Belleau, B. D., and Xu, Y. (2006). Predicting purchase intention of a controversial luxury apparel product. *Journal of Fashion Marketing and Management: An International Journal*, 10(4), 405–419. doi: 10.1108/13612020610701947.

Sun, T., Viswanathan, S., Huang, N., & Zheleva, E. (2021). Designing promotional incentives to embrace social sharing: Evidence from field and online experiments. *MIS Quarterly*, *45*(2), 789–820. https://doi.org/10.25300/misq/2021/15352.

Sun, T., Viswanathan, S., Huang, N., & Zheleva, E. (2021). Designing promotional incentives to embrace social sharing: Evidence from field and online experiments. *MIS Quarterly*, *45*(2), 789–820. https://doi.org/10.25300/misq/2021/15352.

Tajfel, H. (1974). Social identity and intergroup behaviour. *Social Science Information*, *13*(2), 65–93. https://doi.org/10.1177/053901847401300204.

Tajfel, H., and Turner, J. C. (1979). An integrative theory of inter-group conflict. In W. G. Austin and S. Worchel (Eds.), *The social psychology of inter-group relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tajfel, H., Billig, M. G., Bundy, R. P., and Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149-178. doi:10.1002/ejsp.2420010202.

Tärning, B., and Silvervarg, A. (2019). “I Modeln’t Understand, I´m Really Not Very Smart”—How Design of a Digital Tutee’s Self-Efficacy Affects Conversation and Student Behavior in a Digital Math Game. *Education Sciences*, *9*(3), 197. https://doi.org/10.3390/educsci9030197.

Taylor, D. M., and Doria, J. R. (1981). Self-Serving and Group-Serving Bias in Attribution. *The Journal of Social Psychology*, 113(2), 201–211. https://doi.org/10.1080/00224545.1981.9924371.

Taylor, S. E. (2011). Social support: A review. In H. S. Friedman (Ed.), *The Oxford handbook of health psychology* (pp. 189–214). Oxford University Press.

Thomas, J., Noel-Storr, A., Marshall, I., Wallace, B., McDonald, S., Mavergames, C., Glasziou, P., Shemilt, I., Synnot, A., Turner, T., Elliott, J., Agoritsas, T., Hilton, J., Perron, C., Akl, E., Hodder, R., Pestridge, C., Albrecht, L., Horsley, T., … Pearson, L. (2017). Living systematic reviews: 2. Combining human and machine effort. In *Journal of Clinical Epidemiology* (Vol. 91, pp. 31–37). Elsevier USA. https://doi.org/10.1016/j.jclinepi.2017.08.011.

Tubadji, A., Huang, H., and Webber, D. J. (2021). Cultural proximity bias in AI-acceptability: The importance of being human. *Technological Forecasting and Social Change*, *173*, 121100. https://doi.org/10.1016/j.techfore.2021.121100.

Turner, J. C., and Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence.

British Journal of Social Psychology, 25(3), 237–252. https://doi.org/10.1111/j.2044-8309.1986.tb00732.x.

Turner, J. C.; Reynolds, K. H. (2001). Brown, S. L.; Gaertner (eds.). "The Social Identity Perspective in Intergroup Relations: Theories, Themes, and Controversies". Blackwell Handbook of Social Psychology: Intergroup Processes. 3 (1): 133–152. doi:10.1002/9780470693421.ch7.

Turner, J. H. (1988). *A theory of social interaction*. Stanford University Press.

Valle-Cruz, D., and Sandoval-Almazan, R. (2018, May 30). Towards an understanding of Artificial Intelligence in government. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3209281.3209397.

Van den Broeck, E., Zarouali, B., and Poels, K. (2019). Chatbot advertising effectiveness: When does the message get through? *Computers in Human Behavior*, *98*, 150–157. https://doi.org/10.1016/j.chb.2019.04.009.

Van der Maas, H. L. J., Snoek, L., and Stevenson, C. E. (2021). How much intelligence is there in artificial intelligence? A 2020 update. *Intelligence*, *87*. https://doi.org/10.1016/j.intell.2021.101548.

Vanichvasin, P. (2021). Chatbot Development as a Digital Learning Tool to Increase Students' Research Knowledge. International Education Studies, 14(2), 44-53.

Verganti, R., Vendraminelli, L., and Iansiti, M. (2020). Innovation and Design in the Age of Artificial Intelligence. *Journal of Product Innovation Management*, *37*(3), 212–227. https://doi.org/10.1111/jpim.12523.

Vetter, J. (2000). Performance analysis of distributed applications using automatic classification of communication inefficiencies. *Proceedings of the 14th International Conference on Supercomputing - ICS '00*. https://doi.org/10.1145/335231.335255.

Wanberg, C. R., & Kammeyer-Mueller, J. D. (2000). Predictors and outcomes of proactivity in the socialization process. Journal of applied psychology, 85(3), 373.

Wang, D., Churchill, E., Maes, P., Fan, X., Shneiderman, B., Shi, Y., and Wang, Q. (2020). From human-human collaboration to human-ai collaboration. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3334480.3381069.

Webb, M., Autor, D., Bloom, N., Bresnahan, T., Brynjolfsson, E., Chetty, R., Coyle, D., Gentzkow, M., Hoxby, C., Jaravel, X., Jones, C., Klenow, P., Pistaferri, L., Rafey, W., Sorkin, I., Reenen, J. van, Agrawal, A., Connelly, T., Han, A., Thornton, G. (2019). *The Impact of Artificial Intelligence on the Labor Market*. https://web.stanford.edu/.

Weidemann, A., and Rußwinkel, N. (2021). The Role of Frustration in Human–Robot Interaction – What Is Needed for a Successful Collaboration? *Frontiers in Psychology*, 12. https://doi.org/10.3389/fpsyg.2021.640186.

Wilkens, U. (2020). Artificial Intelligence in the workplace – a double-edged sword. *The International Journal of Information and Learning Technology*, *37*(5), 253–265. https://doi.org/10.1108/ijilt-02-2020-0022.

Wirth, N. (2018). Hello marketing, what can artificial intelligence help you with? *International Journal of Market Research*, *60*(5), 435–438. https://doi.org/10.1177/1470785318776841.

Xu, K., and Lombard, M. (2017). Persuasive computing: Feeling peer pressure from multiple computer agents. *Computers in Human Behavior*, *74*, 152–162. https://doi.org/10.1016/j.chb.2017.04.043.

Yang, H., & Oliver, M. B. (2010). Exploring the effects of television viewing on perceived life quality: A combined perspective of material value and upward social comparison. *Mass Communication and Society*, *13*(2), 118–138. https://doi.org/10.1080/15205430903180685.

Yu, C. E. (2019). Humanlike robots as employees in the hotel industry: Thematic content analysis of online reviews. *Journal of Hospitality Marketing and Management*, *29*(1), 22–38. https://doi.org/10.1080/19368623.2019.1592733.

Yu, H., Shen, Z., Miao, C., Leung, C., Lesser, V. R., and Yang, Q. (2018). *Building Ethics into Artificial Intelligence*. http://arxiv.org/abs/1812.02953.