



**Xinyi Yang**

**UNDERSTANDING CHATBOT SERVICE ENCOUNTERS: CONSUMERS'  
SATISFACTORY AND DISSATISFACTORY EXPERIENCES**

Master's Thesis

Marketing

June 2020

Unit Department of Marketing, Management and International Business			
Author Xinyi Yang		Supervisor Eeva-Liisa Oikarinen (Assistant Professor) and Saila Saraniemi (Professor)	
Title Understanding Chatbot Service Encounters: Consumers' satisfactory and dissatisfactory experiences			
Subject Marketing	Type of the degree Master's Degree	Time of publication June 2020	Number of pages 81+2
Abstract			
<p>The service industry keeps growing these years. Artificial intelligence (AI) has started to be used in the service industry gradually, and the service chatbot is an excellent example of this phenomenon. Many giants have applied chatbots to handle their consumer services, such as LATTJO from IKEA, Stylebot from Nike, and Siri from Apple.</p> <p>Understanding the advanced chatbot service experiences can help companies to optimize their chatbot services and improve their consumers' satisfaction, which can bring them positive word-of-mouth, customer loyalty, re-purchase behavior, etc. However, chatbot services is an edge research area with limited studies about it. Thus, having the most advanced understanding of chatbot service experiences becomes particularly important. This study intends to fill this gap from chatbot service encounters' perspective by understanding consumers' satisfactory and unsatisfactory experiences with chatbots.</p> <p>Due to this study focuses on chatbot service encounters and online customer service experiences, a qualitative research method be applied because it enables data to be explainable and justifiable. Data collection methods consist of the critical incident technique (CIT) and the online focus group. In the end, 22 validity incidents were collected.</p> <p>Through data analysis, the author developed an incident sorting process and concluded eight types of chatbot service encounters within three groups by this process. The three groups are chatbot response to after-sales services, chatbot response to consumers' needs, and unprompted chatbot actions. Moreover, 16 sources of different types of chatbot service encounters were found. Based on all the findings stated above, this study created an integrated framework for chatbot service encounters in online customer service experiences.</p> <p>In conclusion, this study develops theoretical contributions by developing the integrated framework, creating an incident sorting process, and finding the sources for different service encounters. Based on these findings, this study also provides some managerial implications that companies could use to manage their chatbot services.</p>			
Keywords Critical Incident Technique (CIT), online customer service experience (OCSE), chatbot service encounter			
Additional information			

## CONTENTS

<b>1</b>	<b>INTRODUCTION.....</b>	<b>6</b>
1.1	Background of the research topic and the research gap .....	6
1.2	The aim of the study and research questions.....	9
1.3	Key concepts .....	10
1.4	Research methodology .....	11
1.5	Structure of this study .....	12
<b>2</b>	<b>CUSTOMER EXPERIENCE – FROM TRADITIONAL TO 2.0 (ONLINE)</b> .....	<b>14</b>
2.1	Service encounters in traditional customer experiences.....	14
2.2	Online customer service experience (OCSE).....	17
2.2.1	Service encounter 2.0.....	18
2.2.2	Online service encounter satisfaction .....	19
<b>3</b>	<b>CHATBOT SERVICE ENCOUNTERS IN OCSES .....</b>	<b>21</b>
3.1	Human-machine communication (HMC) .....	21
3.2	Conceptualization of the chatbot .....	22
3.3	The benefits and barriers of chatbots.....	24
3.4	Chatbot service encounter .....	26
3.5	Integrative framework: Chatbot service encounters in online customer service experiences .....	29
<b>4</b>	<b>METHODOLOGY .....</b>	<b>34</b>
4.1	Combination of CIT and focus group methods.....	34
4.2	Data collection process.....	37
4.3	Data analysis method .....	38
<b>5</b>	<b>FINDINGS OF THE EMPIRICAL ANALYSIS .....</b>	<b>41</b>
5.1	Critical incident classification of chatbot service encounters .....	41
5.1.1	Major groups of chatbot service encounters .....	42

5.1.2	Chatbot service encounter segmentation .....	43
<b>5.2</b>	<b>Sources of satisfactory and dissatisfactory chatbot service encounters .....</b>	<b>49</b>
5.2.1	Sources of satisfactory chatbot service encounters.....	50
5.2.2	Sources of dissatisfactory chatbot service encounters .....	52
<b>5.3</b>	<b>Classifying the dimensions of chatbot service encounters.....</b>	<b>54</b>
<b>5.4</b>	<b>Summary of the empirical findings .....</b>	<b>56</b>
5.4.1	Summary of the satisfactory service encounter .....	56
5.4.2	Summary of the dissatisfactory service encounter.....	57
<b>6</b>	<b>CONCLUSIONS .....</b>	<b>61</b>
<b>6.1</b>	<b>Answers to the research questions.....</b>	<b>61</b>
<b>6.2</b>	<b>Theoretical contributions .....</b>	<b>64</b>
<b>6.3</b>	<b>Managerial implications .....</b>	<b>67</b>
<b>6.4</b>	<b>Evaluations of this study: validity and reliability .....</b>	<b>69</b>
<b>6.5</b>	<b>Limitations of the study .....</b>	<b>70</b>
<b>6.6</b>	<b>Suggestions for future research .....</b>	<b>71</b>
	<b>REFERENCES.....</b>	<b>73</b>
	<b>APPENDICES .....</b>	<b>82</b>
	<b>Appendix 1 PRE-QUESTIONNAIRE.....</b>	<b>82</b>
	<b>Appendix 2 SEMI-STRUCTURED FOCUS GROUP DISCUSSION .....</b>	<b>83</b>

## **FIGURES**

<b>Figure 1. Integrative framework: Chatbot service encounters in OCSEs.....</b>	<b>33</b>
<b>Figure 2. Incident sorting process (in the chatbot situation).....</b>	<b>40</b>
<b>Figure 3. Summary: Chatbot service encounters in OCSEs. ....</b>	<b>62</b>

## **TABLES**

<b>Table 1. Summary of the literature review about chatbot service encounters in OCSEs.....</b>	<b>29</b>
<b>Table 2. The sources of different chatbot service encounters and its dimensions. ....</b>	<b>31</b>
<b>Table 3. Information for the focus group discussions. ....</b>	<b>38</b>
<b>Table 4. Group and category classification by type of incident outcome.....</b>	<b>41</b>
<b>Table 5. Group 1: Chatbot response to after-sales services.....</b>	<b>44</b>
<b>Table 6. Group 2: Chatbot response to consumers' needs. ....</b>	<b>47</b>
<b>Table 7. Group 3: Unprompted chatbot actions.....</b>	<b>48</b>
<b>Table 8. Classifying the dimensions of chatbot service encounter. ....</b>	<b>55</b>

## 1 INTRODUCTION

This Master's thesis focuses on the chatbot service encounter through understanding consumer satisfactory and dissatisfactory experiences. In this chapter, the research background and gap are presented in the beginning. Next, research questions are introduced and following the key concepts. Then, this chapter offers a brief overview of the research method. The overall structure of this study demonstrates briefly at the end of this chapter.

### 1.1 Background of the research topic and the research gap

Nowadays, 70% of global GDP is contributed by the service industry and is expected to keep growing. One stimulus behind this phenomenon is the advanced digital technologies (Wan & Chan, 2019). Consumers are consuming more time and money online for both physical products and services, which create more "online participation" (Lee & Lee, 2020). In the digital world, an increasing number of users are using artificial intelligence (AI) to assist their businesses (Devaney, 2018). The robot is one of the products of AI that has started to be used in the service industry gradually, and the chatbot is an example of this kind of service robots.

Chatbots were predicted to take care of 85% of online customer service interaction by 2020 (Julia, 2018). The most common channel to interact with chatbots include text messages, individual apps, and Messenger (from Facebook) (Dal Porto, 2017). Many giants have applied chatbots to handle their consumer services, such as Microsoft, Google, IBM, etc. (Ranjan & Mulakaluri, 2018). This study focuses on utilitarian text-based chatbots because it is getting popular progressively among online services with its benefits for both consumers and companies.

The previous studies have stated that chatbots allow companies to offer consumers continuous services faster and more efficiently and help companies save costs by saving human resources (Dal Porto, 2017). For example, machines can identify consumers' emotions through algorithms during the service process. This kind of identification provides chatbots potential capacities to serve consumers better than human employees. (Huang & Rust, 2018.) Besides, consumers can benefit from the

chatbot services too. For example, chatbots allow consumers to access services anytime and anywhere in a productive (ease, speed, and convenience) way (Devaney, 2018; Brandtzaeg & Folstad, 2017. Wunderlich & Paluch, 2018). Many companies are trying to explore how to enhance and optimize their chatbot services to please their consumers. This kind of exploration increases the importance of understanding consumers' experiences with chatbot services, which makes this study more valuable.

Different from the traditional service experiences, the alter of service providers (from humans to machines) has changed consumer behaviors. Consumers interact with salespersons directly in the traditional human service context. Salespersons play an essential role because consumers prefer to purchase products/services from the salespersons they are familiarized (Trotter, 2017). However, in the chatbot context, the situation is different because consumers interact with a string of emotionless codes instead of emotional salespersons. For example, customers can directly chat with the chatbots or use tablets to finish their orders in a restaurant without waiters (Garber, 2014). This kind of role change has changed the service encounters.

The service encounter is an ongoing exchange of value (Kleinschafer, Morrison & Dowell, 2018), which directly affects customer satisfaction and further affects customer loyalty, word-of-mouth, and re-purchase decision, etc. The service encounter has been stated as further essential in the digital world, because of the Internet speeds up many things, such as the spread of negative word-of-mouth, easy access to negative comments, etc. (Cyr, 2008). Thus, it is necessary and essential to have an advanced understanding of the chatbot service encounters.

The above contents about chatbots, chatbot service encounters, and consumer behaviors indicate the importance of having an advanced understanding of chatbot service encounters in online consumer service experiences. Nevertheless, the chatbot service is an edge research area with many research gaps and calls for more studies.

First and foremost, there are limited amount of studies that focus on chatbot service encounters. The service encounter is a crucial topic for businesses and was discussed by many researchers already (e.g., Surprenant & Solomon, 1987; Bitner, Booms & Tetreault, 1990; Larivière et al., 2017). However, only a small number of existing

studies focus on the chatbot service encounter (e.g., Mimoun, Poncin & Garnier, 2012; Feine, Morana & Gnewuch, 2019; Wunderlich & Paluch, 2018). Chatbot service encounters are different from the traditional service encounters to a certain degree. Specifically, consumers are involved in the self-service process when interacting with chatbots (Huang & Rust, 2018). This kind of self-services changed consumer behaviors because the service provider changed from humans to machines. Individuals behave differently when facing different situations and communicators (Mou & Xu, 2017). For example, individual's responses to greetings from chatbots are slower than responses to humans (Kanda et al., 2008). However, this situation is not absolute. However, some studies have demonstrated that people are similarly responding to virtual agents/chatbots compared to humans if they perceive chatbots have human characteristics (such as friendliness) (Verhagen et al., 2014). In other words, we humans are likely to see chatbots' characteristics as humanlike due to anthropomorphism (Lee, 2018).

Second, there is a gap in research methods. Some previous studies were discussed about chatbot service encounters by different methodologies. Mimoun, Poncin and Garnier's (2012) research have used in-depth interviews, which allowed them to catch descriptive data about service encounters. Nonetheless, that research is from 8 years ago, and AI has made massive progress during the most recent years, which means the existing conclusions are not convincing anymore. Wunderlich and Paluch (2018) have applied the think-aloud and purposive sampling methods in their study. Participants were asked to finish some tasks, and then they have to present their mind after completing these tasks. In this way, the author can observe participant behaviors, but they do not have experience using this method, which means that their study has some limitations. Feine, Morana and Gnewuch (2019) have applied sentiment analysis and automated methods to analyze chatbot service encounter satisfaction through analyzing users' written text. Their study used the old dialog corpus, which cannot ensure data's validity. The critical incident technique (CIT) is an appropriate method for the customer experience studies, particularly to understand the service encounters (Bitner, Booms & Tetreault, 1990). However, it has not been used in the chatbot service encounter study. Thus, the author applies CIT together with the focus group method in this study to understand chatbot service encounters by analyzing consumers' satisfactory and dissatisfactory experiences with chatbots.



Additionally, the author has a personal interest in "AI in business". As a member of the young generation, it is unavoidable to use chatbot services in daily life. However, the unpleasant experiences with chatbot services are always happening. Thus, the author would like to acquire more knowledge about the chatbot services by understanding young consumers' experiences with chatbots in this thesis.

## **1.2 The aim of the study and research questions**

This Master's thesis aims to find out the advanced framework of chatbot service encounters in online service experiences through understanding consumers' satisfactory and dissatisfactory experiences with chatbots. Then, the study intends to make theoretical contributions to the existing literature on online service experiences and chatbot service encounters. The study also provides some suggestions for companies to optimize their chatbot services. This study focuses on the utilitarian text-based chatbot services, such as question consulting, popped up services, etc. Moreover, the study is focusing on consumers (people who consume the products) instead of customers (people who purchase the products). More arguments for these choices are explained in the next section. Based on those mentioned above, the main research question for this study is:

*What is the theoretical framework of the chatbot service encounters in online customer service experiences?*

Additionally, three supporting questions for this study are identified:

- *What is the incident sorting process for the chatbot service encounter?*
- *What are the sources of satisfactory and dissatisfactory chatbot service encounters?*
- *What are the dimensions of chatbot service encounters?*

### 1.3 Key concepts

In this part, the main terms and concepts associated with this study are explained based on the existing literature. First and foremost, this section discusses and provides the definitions of *online customer service experience*, *service encounter 2.0 (online service encounter)*, *customer satisfaction*, and *chatbots*. Then, the uses of concepts of *customer* and *consumer* in this study are explained to avoid confusion.

Customer experiences are shifting towards the digital consumer experiences. In the beginning, consumer experiences were used to focus on offline services. Later in 2013, Klaus (p. 448) developed the concept of *online customer service experience (OCSE)*. The chatbot service experiences can be seen as part of the online service experiences. Therefore, the concept of Klaus will be used in this study. Besides, OCSE is a part of the online customer experience (OCE). Thus, the judgments and discussions in Chapter 2 and Chapter 3 of OCE are reasonably suitable to the OCSE.

*Online customer service experience (OCSE): “the customers’ overall mental perception of their interactions with the online service provider and other customers expressed in its dimensions functionality and psychological factors”.*

Service encounters were perceived earlier as the interaction between the service providers and consumers. Lately, more elements are involved into this concept, such as the environment, technology, network, etc. (Patrício, et al., 2011; Tax, McCutcheon, & Wilkinson, 2013). However, there is no existing definition for the chatbot service encounter so far. Therefore, this study applies the definition of *service encounter 2.0 (online service encounter)* from Larivière et al. (2017, p. 2) due to chatbot services are a part of the online services.

*Service encounter 2.0: “any customer-company interaction that results from a service system that is comprised of interrelated technologies (either company- or customer-owned), human actors (employees and customers), physical/digital environments and company/customer processes”.*

With the same reason as the chatbot service encounters, this study applied the *customer satisfaction* concept from Oliver (1981, p. 27), which can be seen as the foundation of the chatbot service encounter satisfaction.

*Customer satisfaction: "It means product/service performance perceived by consumers higher than their expectations".*

Service bots were named in different ways, the most common one was the *virtual agents*, which was adopted by many studies (e.g., Brave & Nass, 2002; Groom et al., 2009). Nowadays, "*chatbot* has become the mainstream word. This study focused on the utilitarian text-based chatbots and applied the definition from Dale (2016, p. 813).

*Chatbot: "any software application that engages in a dialog with a human using natural language".*

Last but not least, it is necessary to make the difference between the "*customer*" and "*consumer*" to avoid confusion. *Customer* refers to people who paid for products and services. *Consumer* refers to people who used products and services (Anonymous, 2001, p. 101), it is more from users' perspective. Thus, employees ask services from companies' chatbots also be counted as consumers. This study adopts the *consumer* because it is focusing on the utilitarian text-based chatbots, which also provide services for employees. However, in the introduction, theoretical framework, and conclusion chapters, both words are appearing because the *customer* as terminology was used in the previous researches widely, such as customer satisfaction, customer loyalty, etc. (e.g., Oliver, 1981; Cronin & Taylor, 1992; Noone et al., 2009).

#### **1.4 Research methodology**

Qualitative studies can provide a comprehensive and contextual understanding of consumer experiences (Polit & Beck, 2010). It also enables data to be more explainable and justifiable, which can help researchers to understand a phenomenon better (Diekroger, 2014). The previous part has mentioned that this study is surrounding the chatbot service encounters, which is a part of OCEs. Thus, a qualitative research method is used in this study with the critical incident technique (CIT) & focus group (online) approaches combo as the data collection method.

The CIT can help researchers to find out high-quality information from participants' satisfactory and dissatisfactory experiences (Viergever, 2019). Bitner, Booms and Tetreault (1990) have proved this is an appropriate way for the study of service

encounters. The focus group is a suitable way to collect information about individuals' experiences (Hines, 2000). This method allows researchers to involve in the group discussion to collect rich data with lower costs.

In this study, both methods are applied. The focus group is a ministrant method that helps the author to have a deeper understanding of the critical incidents. Furthermore, it is necessary to mention that this study also used pre-questionnaires before the focus group discussion, which intends to guarantee all the participants' experiences able to match the requirements of the critical incidents. At the same time, it also can help participants to comprehend the research topic better.

Besides, this study applied the abductive strategy, which focuses on the "meanings and interpretations, the motivations and intentions" in people's daily life. It means describing and understanding the social life, such as people's actions and nature of objects. The logic of abductive strategy is from lay concepts (general formulation of the problem), then to generate ideal types, finally to develop an interpretation or construct a theory. (Ong, 2012.) Reflecting on this study, the author started from diagnosing unpleasant experiences with chatbot services and reading existing studies (generate the research idea), then the critical incidents were collected by pre-questionnaires and focus group discussions. Finally, the framework of the chatbot service encounters in OCSEs was developed based on the theoretical study and empirical findings.

## **1.5 Structure of this study**

This study consists of six chapters that covered both theoretical chapters and empirical study. The research questions are the "beacon light" for all chapters in this study, and each chapter's main contents are presented briefly in this section.

Chapter 1 provides a blueprint for this study by introducing the background information and the main research idea to readers. Next, Chapter 2 and Chapter 3 aim to demonstrate the existing studies about OCSEs and chatbots service encounters by discussing and evaluating the relevant literature. Each chapter consists of a few sub-chapters, such as OCEs, chatbot service encounters, etc. At the end of Chapter 3, the

author summarized some critical concepts with its descriptions in Table 1 and structured them to Figure 1 (chatbot service encounters in the OCSE), which could be seen as the foundation for the answer of the main research question. Then, the object of Chapter 4 is to introduce the data collection process and data analysis methods. The CIT and focus group are presented first and following the data collection process. The incident sorting process developed by this study is presented at the end of this section. Chapter 5 discusses the empirical findings of this study according to the research questions, which is also discussed with the existing studies (similarities and differences). Chapter 6 intends to conclude the entire study. It provides insights into the research questions first and following by the general overview of this study's contributions from theoretical and managerial perspectives. Then, the study's evaluations and limitations are examined, which are related to the theoretical and methodological aspects. The suggestions for future studies are placed in the end.

## **2 CUSTOMER EXPERIENCE – FROM TRADITIONAL TO 2.0 (ONLINE)**

This chapter demonstrates the existing literature about customer experiences, and mainly focuses on the OCE and OCSE. The first sub-chapter introduces traditional customer experience literature. Due to this study focuses on the OCSE, the purpose of traditional customer experiences is to set the stage for the OCE because it is the origin of the OCE. In the second sub-chapter, the OCSE and its differences with the traditional customer experiences are introduced first, following the discussion of the service encounter 2.0 and online service encounter satisfaction. This sub-section intends to pave the way for the next chapter, because of chatbot services are affiliation to online services.

### **2.1 Service encounters in traditional customer experiences**

Klaus (2020) stated that customer experience plays the “iron throne,” as it is the only thing that can be managed by the companies. Carbonne and Haeckel (1994, p. 9) defined the customer experience as “*the take-away impression formed by people’s encounter with products, services, and businesses*”. In this situation, service encounters are formed by service employees and customers, which makes individual’s emotion plays an influential role as the employees face consumers directly (Skowron, 2010), and customers draw upon this kind of encounter (the service they received) to evaluate the service quality (Gupta & Zeithaml, 2006).

Encounters, shop atmosphere, facilities, post-transaction services, etc. converged into the customer experience (Resnick, Foster & Woodall, 2014). This study mainly focuses on service encounters instead of other elements. The traditional service encounter means “*dyadic interaction between a customer and service provider*” (Surprenant & Solomon, 1987, p. 87). It is from customers’ point of view to talk about the interaction between customers and companies. (Surprenant & Solomon, 1987; Bitner, Booms & Tetreault, 1990).

One of the most important studies of service encounters was from Bitner, Booms and Tetreault (1990). Their study was focused on three services-oriented industries – hotels, restaurants, and airlines. They applied the CIT as the data collection method, and about

700 validated incidents about satisfactory and dissatisfactory service encounters were collected in total. Then, they developed an incident sorting process to analyze their data. The whole process is like a "decision tree" with branches and leaves in a flowchart-like structure. It consists of three main branches, and each of them represents one question. These questions surround three attributes: services (itself), needs, and employee actions. Leaves stand for each question's outcomes, and each leaf stands for a label (one category). The process starts with a simple question "is there a service delivery system failure". Answering yes means the service failure, then the process goes to the first branch: nature of service failures. In contrast, answering no goes to the next question "is there an implicit/explicit request for accommodation". Replying yes goes to the second branch: nature of requests/needs. Answering no goes to the third question, "is there and unprompted/unsolicited action by employees". Responding yes goes to the third branch: the nature of employee actions. In summary, the whole process starts from the main category and then goes into small categories step by step. This "decision tree" logic helped Bitner, Booms and Tetreault found 12 types of service encounters with human employees. It inspires the author to apply this logic in this current study, and more details are presents in Chapter 4.

The extant literature about the service encounter suggests that service employees directly affect service quality due to their emotions can influence customers' emotions (Resnick, Foster & Woodall, 2014). For example, there are two types of behavior according to different service encounters – citizenship- and dysfunctional behaviors. Citizenship behavior means favorable behaviors, such as employees' voluntary behaviors with positive effects. This kind of behavior is able to encourage customers' citizenship behavior and generate customer satisfaction. In contrast, dysfunctional behavior means unfavorable behaviors with negative effects (customer dissatisfaction). (Yi & Gong, 2008) All in all, different service encounters result in different customer attitudes (satisfaction or dissatisfaction), which affects customer overall experiences.

Customer satisfaction is the outcome of customer experiences, which can cause customers to generate emotional reactions towards products/services (Oliver, 1981). Emotional reactions are based on the gap between customer expectations and product/service performances received by customers (Tse & Wilton, 1988). Based on this kind of gap, customers can generate an overall feeling towards a company (Cronin

& Taylor, 1992). If the product/service performances perceived by customers higher than their expectations, customer satisfaction will be generated.

Customer satisfaction is a cumulative judgment affected by consumer's post-purchase experiences (Van Doorn & Verhoef, 2008). There is no denying that customer satisfaction is essential for every company because it is an influential factor to increase companies' turnover and revenue (Noone et al., 2009). Especially for e-commerce companies (Cyr, 2008), due to the Internet speed up the information spread process. Besides, customer satisfaction positively affects customer loyalty, word-of-mouth, companies' profits, and favorable purchase intentions (Reynolds & Beatty, 1999; Bowen & Chen, 2001).

In contrast, customer dissatisfaction happens when product or service performances perceived by customers is lower than their expectations. The service failure is the main reason for customer dissatisfaction. It has many situations that can cause service failures. For example, in the traditional service process, burnout attitudes from employees (service providers) have negative impacts on customer satisfaction. The reasons that cause employees' burnout attitudes include poor salaries, poorly understanding from managers or consumers, and consumer abuse. These elements can bring employees physical and mental problems. (Söderlund, 2017.)

One essential method to change customer dissatisfaction to satisfaction is to provide efficient service recoveries. Service recoveries cannot work efficiently without understanding customer satisfaction, which can turn customer dissatisfaction to high levels of satisfaction. This kind of transaction can generate positive word-of-mouth and future repurchase attentions (Bitner, Booms & Tetreault, 1990; Halstead & Page, 1992; Smith, Bolton, & Wagner, 1999; Wallin Andreassen, 2000; Maxham & Netemeyer, 2002). Thus, it calls the need for companies to understand customer satisfaction and dissatisfaction.

However, these kinds of situations might do not exist when people are communicating with a machine for many reasons. In the next section, the literature of OCSE is demonstrated.



## 2.2 Online customer service experience (OCSE)

In the digital environment, the situation for customer experiences is more dynamic compared to traditional customer experiences (Klaus, 2013). The definition of *online customer experience (OCE)* from Trevinal and Stenger (2014, p. 324) is “*a holistic and subjective process resulting from interactions between consumers, shopping practices, and the online environment.*” The concept of OCE emphasizes online customer-organization interactions, which could be information searching, purchasing products, using services, etc. This study mainly focuses on online services. Thus, a more detailed definition towards online services – *online customer service experience (OCSE)* from Klaus (2013, p. 448) will be applied: “*the customers' overall mental perception of their interactions with the online service provider and other customers expressed in its dimension's functionality and psychological factors*”. This definition is related to customers' mental perception which matches with the idea of this study about collecting critical incidents from consumers because the critical incidents can reflect consumers' mental perception about their attitudes and experiences with chatbot services.

The online service is a kind of untact (un-contact) service (Lee & Lee, 2020); it means service providers and consumers are not necessary to have face-to-face interactions. Therefore, the online service context is different from the offline context. The online context has lower personal contacts, intensive information provision, consumer dictations for the interactions (anytime and anywhere), and audio-visual brand presentation. In specific, 1) the offline environment provides more face-to-face interaction than the online environment, 2) the online environment is able to bring consumers more information than offline (poster, brochures, etc.) 3) the online services can happen anytime & everywhere. However, offline services are always oriented by organizations, 4) the brand presentation affected by the employees and tangible devices in the offline environment. In the online environment, the brand presentations are always in an audio-visual way. (Rose, Hair & Clark, 2011)

An online customer-organization interaction is formed by both cognitive (goal-oriented and rational) and affective (emotional) information processing (Rose, Hair & Clark, 2011). It means that the quality of OCSE received by consumers is related to

both rational and emotional factors (Chaffey & Ellis-Chadwick, 2016). The factors could be the quality of the website (the website performance), online consumer behavior (such as how consumers search the information online), and industries (Rose, Hair & Clark). This study does not care about the website and the industry factors but mainly focuses on online consumer behavior.

Online consumer behaviors are different from offline consumer behaviors due to customers are playing different roles. In the online context, customers could be visitors, users, etc. (Cho & Park, 2014). Simultaneously, the service provider could be humans, machines (like chatbot), etc. Machines can speed up the service response time and improve e-service efficiency (Li, 2014), and Chapter 3 will present more information about how machines are used in businesses.

### 2.2.1 Service encounter 2.0

The service encounter is changing along with the development of technologies. Comparing with the traditional service encounter, service encounter 2.0 involves more “players”, such as the environments and technologies. Larivière et al. (2017, p. 2) defined it as “*any customer-company interaction that results from a service system that is comprised of interrelated technologies (either company- or customer-owned), human actors (employees and customers), physical/digital environments and company/customer processes*”. It is about the complexity of interactions between humans and technologies which is match with this study (consumers and chatbots). Under this definition, both employees and consumers are playing different roles compare to the traditional service encounters. This kind of difference caused the service encounter 2.0 is distinct from traditional service encounters and made this study more necessary.

For the service encounter 2.0, on the one hand, the employee plays four types of roles – enabler, innovator, coordinator, or differentiator. These various types of roles indicate that human employees and technologies are supporting each other and working together. From enablers' perspective, the role of employees is like the bridge between consumers and techniques and ensure they can play their own roles well. However, if this bridge did not handle the situation well, it may lead to adverse

outcomes. From innovators' perspective, employees can help companies to find actively pinpoint areas for service improvement through detecting consumer needs. From coordinators' perspective, multi-channel can provide consumers different experiences, but it requires employees to optimize outcomes from different service encounters. From differentiators' perspective, employees have some particular service skills which are less replicable by machines, such as machines do not have feelings like humans. (Larivière et al., 2017; Bowen, 2016.) In a word, the relationship between humans and machines is like a partnership, and these two parties working together can have better performances.

On the other hand, the consumer plays as "partial employees" in the online context, which means they act as co-creators of the service encounter (Mills, Chase & Margulies, 1983; Bowen, 1986; Larsson & Bowen, 1989; Prahalad & Ramaswamy, 2004; Larivière et al. 2017). For example, consumers can help companies optimize their services by sharing their personal information because companies can know them better in this way (Chan, Yim & Lam, 2010). In other word, if the machine as the service provider, the dialog corpus used to store the conversations with consumers is a valuable information source for companies to know their consumers.

### 2.2.2 Online service encounter satisfaction

2.1.2 section mentioned that customer experiences are able to cause customers' emotional reactions towards products or services, i.e., customer satisfaction. In the digital world, customer satisfaction is the result of positive cognitions of OCE (Rose, Hair & Clark, 2011). It is the same as the traditional service experiences; if customers perceive products or services' performances as higher than their expectations (positive cognition), customer satisfaction will be generated. This study focuses on customer satisfaction in services. The quality of online customer services affects online customer satisfaction (Wu et al., 2012), and the service quality is a crucial feature for consumers to evaluate an e-commerce company (Li, 2014).

This study mainly focuses on chatbot service encounters. Thus, this part concentrates on the service encounter satisfaction instead of discussing service satisfaction in general. The relationship between service encounters and service encounter

satisfaction is the causal relationship. Service encounter satisfaction is the measure of consumers' satisfaction in transactions. The traditional service encounter satisfaction has strong impacts on consumers' overall satisfaction for the whole service experiences (Verhagen et al., 2014; Caruana, 2002.), and it is similar to the online situation. The online encounter satisfaction positively affects consumers' overall satisfaction toward companies (Chan, Barnes & Fukukawa, 2016).

Online service encounter satisfaction is affected by many factors. Wolfinbarger and Gilly (2003) stated that information comprehensiveness and service process efficiency are two factors that influence the service encounter satisfaction. Understandably, smooth service processes are always able to please their consumers. Koufteros, Verghese, and Lucianetti (2014) noted that the delivery of information plays an essential role in the service encounter satisfaction. Organizations can use the proper information to enhance their capabilities, such as understanding their consumers' expectations and needs. Verhagen et al. (2014) mentioned that service providers' friendliness (polite, responsive, etc.) and professionalization (the capability to provide knowledgeable answers) have substantial effects on service encounter satisfaction. Without denying that in most of the situations, knowledgeable answers are able to meet consumers' expectations.

Companies should set up continuous satisfaction as a part of their strategies (Chan, Barnes & Fukukawa, 2016), because of the service encounter satisfaction can generate positive word-of-mouth, customer loyalty, and repurchase behavior (Oliver, 1997). There are different ways to produce satisfactory service encounters. For instance, companies can try to provide customized and flexible services, handle service failures properly, and reduce the gaps between their service qualities and their consumers' expectations, etc. (Bitner, Brown & Meuter, 2000).

### **3 CHATBOT SERVICE ENCOUNTERS IN OCSES**

This chapter mainly demonstrates the literature which is relevant to chatbot services and online customer service experiences. The first sub-chapter is about human-machine communication (HMC), which paves the way for chatbot services. The second sub-chapter is about the definition of chatbots, and it also points out opportunities for chatbots in businesses. The following section is surrounding the benefits and barriers of chatbots from two perspectives (users and companies). The benefits indicate why companies and consumers should use chatbot services, and the barriers present the challenges for companies and consumers to use the chatbot services. This sub-chapter also stated potential reasons which caused different types of chatbot service encounters. Thus, it can be seen as the transitional phase for the next sub-chapter, which is about chatbot service encounters. It consists of the meaning of chatbot service encounters, differences between chatbot service encounters and the traditional service encounters, consumers' expectations about chatbot services, consumer satisfaction in chatbot service encounters, and how companies should manage their chatbot services. In the last sub-chapter, the author concludes that some core points surround chatbot service encounters and OCSE in Table 1. The sources of satisfactory and dissatisfactory chatbot service encounters are listed separately in Table 2 with three dimensions. Based on these two tables, the author concluded a framework (Figure 1) about chatbot service encounter in online customer service experiences. This framework covers the relationships between different concepts, illustrates the existing consequences of chatbot service encounters, and leaves a place for this study's findings.

#### **3.1 Human-machine communication (HMC)**

Communication has been understood as a social process earlier (Mead, 1967). The communication discipline used to focus on human-human communication, such as an individual expressing information to another individual (Craig, 1999). The human-human context is more "extroverted, conscientious, and self-disclosing" (Mou & Xu, 2017, p. 437). In this context, emotion plays an important role, and it can be viewed as a mediator between consumers and service providers. For instance, service delay may cause consumers' anger emotions, but if the delay time is filled by something

else, the anger emotions can be offset. (Taylor, 1994.) Smile from service providers might bring consumers satisfactory emotions. Thus, it is always necessary to pay attention to the personal relationship between consumers and service providers in the human-human context, as it affects consumer satisfaction. The previous study has stated that the relationship is one of the most critical goals in human-human communications (Hobbs & Evans, 1980). Nevertheless, with the development of technologies, the way of communication has gradually turned from human-human communication to HMC.

The HMC means exchanging information between humans and machines in a clear and precise language. The language in the HMC field means computer programming languages been a very long time, such as C, C+, R, etc. The starting point of HMC can be traced back to 1950 with the question came up by Turing (1950, p. 433) "*Can machines think*". The HMC is developing together with technologies (Rainie & Anderson, 2017). In this context, the question transferred from "*Who is the person interacting with*" to "*What are they communicating with*". As an emerging area of communication, HMC has become a specific research topic. (Guzman & Lewis, 2020.)

In the HMC context, the role of machines has turned from channels to communicators (Guzman & Lewis, 2020). Machines could divide labors, support humans, and enhance humans (Huang & Rust, 2018). Chatbots are an excellent example of this kind of machine communicators.

### **3.2 Conceptualization of the chatbot**

Turing Test in 1950 opened the door for AI, which came up with an idea called "learning machines". (Turing, 1950) Machine learning (ML) means a computer program uses data to improve itself automatically (Mitchell, 1997, p. 2). It is an application of AI and widely used as a supportive technology for consumer services. The chatbot is an excellent example of this kind of technology (Ciechanowski et al., 2018). ML allows chatbots to improve themselves automatically when they are communicating with consumers. It means chatbots are able to adjust the way to interact

with a human by themselves through the “self-learning” process (Guzman & Lewis, 2020).

Chatbots can use human languages to interact with consumers in the virtual conversational services (Lee, Oh & Choi, 2017; Zhang et al., 2017). Using human languages in virtual conversations between humans and machines have become a trend (Sandry, 2018). The natural language processing (NLP) is a supporting tool for it, which helps chatbots to understand and interact with humans in an anthropomorphic way (Devaney, 2018). Chatbots can be found in many industries nowadays, such as banks, health care, airlines, etc. (Feine, Morana & Gnewuch, 2019). Specifically, giant companies can build their service chatbot by themselves, such as LATTJO from IKEA, Stylebot from Nike, Siri from Apple. Small and medium-sized enterprises (SMEs) can develop chatbots through third parties. For instance, over 1,3 billion Facebook Messenger users sent about 20 billion messages among individuals and businesses per month in 2019. This platform has 300,000 Messenger bots built by companies to assist their businesses (Hutchinson, 2019).

The definition of chatbots is “*a text-based or voice-based program which able to mimics the human interactions*” (Ranjan & Mulakaluri, 2018, p.1). One of the earliest chatbots is from the 1960s, and its name was Eliza. Eliza used the simple template-based text to imitate the conversation of a non-directional psychotherapist (Dale, 2016). The definition in 2016 from Dale of the chatbot (p. 813) is “*any software application that engages in a dialog with a human using natural language*”. Later, in Devaney's (2018) report, the chatbot is defined as “*a computer program designed to design to simulate conversation with humans, especially over the Internet*”. This definition covered the definition from Ranjan & Mulakaluri and Dale's study. The first chapter has mentioned that this study focuses on utilitarian text-based chatbots, and the definition from Devaney is closer to this type of chatbot than others. Therefore, in this study, the author applied the definition from Devaney, and the discussions in this chapter are primarily focusing on utilitarian text-based chatbots. Besides, it is necessary to mention that this study also noted the embodied/virtual agents in the text because Dale stated that the chat/conversational agents whose service is based on the text-based interface always means chatbots.

The properties of chatbots provide them a lot of business opportunities. Chatbots are able to be developed on different channels to communicate with consumers by human languages, such as websites (always show up as a chat window), messaging platforms (like Facebook Messenger, WeChat, etc.), and social media (like Twitter, Facebook, etc.). Also, users can use it on different devices (like PCs, mobile phones, etc.). (McTear, Callejas & Griol, 2016.) Besides, the ML allows chatbots are able to improve themselves automatically by “self-learning” processes based on the data collected by themselves or imported by humans (Guzman & Lewis, 2020). Thus, chatbots have been applied in almost all industries to support companies' customer services.

### **3.3 The benefits and barriers of chatbots**

Both benefits and barriers are consisting of two dimensions: users and companies. From the users' perspective, there are some main benefits like services everywhere, anytime (24/7), ease of use, and convenience compared to the human-based services (Brandtzaeg & Folstad, 2017; Devaney, 2018; Wunderlich & Paluch, 2018). Chatbots can automatically help users finish their tasks by conversations, which makes many young people prefer chatbot services because they can get cost-effective solutions (Arcand, 2017).

From companies' perspective, chatbots are able to help companies reduce costs, such as human capital, which allows companies to invest more money in other fields. Moreover, chatbots have less incremental expenses attached to the usage (Wirtz et al., 2018) because of the “self-learning” process. Furthermore, chatbots could help companies filter their service encounters, which means chatbots can handle the conversation first, and if the problem cannot be solved, then hand over to the human employees. In this way, the workload for service employees is partly relieved (Feine et al., 2019). In summary, chatbots could be seen as the proper candidate for traditional customer services (Brandtzaeg & Folstad, 2017). Especially in situations when machines perform better than human labors (Huang & Rust, 2018).

Many companies have proved the benefits of chatbots mentioned above. For example, “Nina” is a chatbot from Swedbank, and the report shows “Nina” can take care of 40,000 conversations in one month, and 81% of the questions were answered



correctly. "Roxy" is a chatbot from the Department of Human Services (DHS) (Australia). It has helped employees to handle 78% of the questions successfully. (Ranjan & Mulakaluri, 2018.)

However, barriers are always existing. On the one hand, Devaney's (2018) report shows that many people are not well prepared to use chatbot services. For individuals who participated in Devaney's interview, 43% of them prefer a real-life assistant. The reason behind this phenomenon might be human employees can: present empathies, identify the subtle linguistic cues, and handle more complex situations (Feine, Moorana & Gnewuch, 2019). Besides, 30% of interviewees are afraid of that chatbots will make mistakes, and 24% of interviewees think chatbots might respond in improper manners. It means users do not know if chatbots can handle their specific needs, which results in many people not finding clear benefits to communicate with chatbots instead of real humans (Arcand, 2017).

On the other hand, there are some challenges faced by companies to build their chatbots. The first challenge is that chatbots are lacking of dialogue data, which could be seen as the foundation for chatbots to training themselves by ML. The training processes always require both quality and quantity dialogue data. The second challenge is that chatbots have poor performances for multi-turn conversations, which is more like a technical challenge. Chatbots in most of the situations have capabilities to deal with the single-turn situations well but not for the multi-turn situations. The "Dialogue Manager" model is a potential solution for this challenge, which allows chatbots to handle the multi-turn situations by using the "self-matching attention" (filter the redundant information) and "sequential utterance-response matching" technologies. However, it requires a lot of resources from companies to develop this kind of system. (Zhu et al., 2018.) The third challenge is that chatbots cannot justify their behaviors similar way as humans. For example, Tay is a chatbot from Microsoft. It was launched on Twitter in 2016, but Microsoft shut it down in 16 hours. The reason is that Tay has learned how to use swearing words, make racist remarks, and inflammatory political statements in these 16 hours (Wakefield, 2016).

### 3.4 Chatbot service encounter

Chatbot service encounters are generated in the context of the HMC based on AI. It is different from traditional service encounters and service encounter 2.0, which has been presented in the second chapter. The chatbot service is part of the online services. However, the chatbot service encounter does not have a theoretical definition so far. Thus, the author applied the definition from Larivière et al. (2017, p. 2) about service encounter 2.0 and the concept of OCSE from Klaus (2013, p. 448), which has been introduced in the previous part to the chatbot service experiences. Both of them have slight differences in the chatbot situation. First, these two concepts have been applied to the term of customers. As mentioned before, this study focuses on consumers instead of customers, because chatbot services are not just for customers, but also for employees and other people who consume chatbot services. Second, human employees are no longer necessary, which means chatbot service encounters are different from the traditional situation. The main reason is that chatbots replaced human employees, but they do not have emotions like humans. The only thing chatbots can do is to read others' feelings and express their feelings by surface (surface-acted emotions) (Wirtz et al., 2018). In a word, consumers can express their feelings to chatbots through text-based messages, but chatbots cannot catch this kind of emotional expression. Last but not least, the OCSE concept is more focused on online purchasing experiences, but this study is concentrated on the chatbot service experiences.

To summarize what has been mentioned above, chatbot service encounters in this study means consumers' perception of interactions with chatbots, which causes consumer satisfaction or dissatisfaction. Understanding chatbot service encounters is explicitly helpful in understanding consumer experiences. Good OCEs can result in positive online behavior, i.e., customer satisfaction (Shobeiri, Mazaheri & Lauoche, 2018).

As digital employees, chatbots' performances in chatbot service encounters are able to generate consumer satisfaction by meeting consumers' expectations. The chatbot service encounter satisfaction in this study means consumer satisfaction in service processes under the interactions between chatbots and consumers. With the same logic of the traditional customer satisfaction presented in the last chapter, the chatbot service

encounter satisfaction will be generated when consumers perceived services performance higher than their expectations. The service quality is always seen as a measurable item in OCSEs, which determines if the services can exceed consumers' expectations. Seck and Philippe (2013) have developed a model of virtual service quality. It covered security, ease of use, information quality, and site design. This model affirmed these factors affect virtual service quality and then positively affect customer satisfaction.

The majority of expectations from consumers side to chatbots consist of (descending sort) 1) providing 24-hour services, 2) getting instant responses, 3) answering simple questions, 4) easy communication processes, 5) solving complaints quickly, 6) good experiences, 7) providing detailed/expert answers, 8) answering complex questions, 9) behaving friendliness and approachability. (Devaney, 2018.) The order of these expectations matches with Arcand's (2017) study that consumers are more ready for chatbots to handle some simple interactions (straightforward information with low knowledge base). However, the expectations listed above are in general situations. Consumers have different expectations for chatbot service encounters among different industries. For example, consumers care more about if the information delivered by chatbots are credible instead of saving time in the luxury industry. In this situation, chatbots should focus on professionalized answers to improve consumer satisfaction. (Chung et al., 2018.)

Except for understanding consumer satisfaction with chatbot service encounters, it is also essential to understand consumer dissatisfaction with chatbot service encounters because companies can revise their service failure based on this kind of understanding. There are many reasons for the dissatisfactory chatbot service encounters. Firstly, lacking online interaction has been a problem for retailers and customers. Therefore, some companies started to use virtual agents on their websites to interact with their customers. However, many virtual agents disappeared after a few years (the data is from France) because there was a gap between customer expectations and customer perceptions for virtual services. Secondly, lacking the intelligence of embodied agents has been another problem. This means that chatbots cannot manage all the information; they fail to understand customers, behave aggressively, and have uncomfortable interaction processes with customers. Thirdly, some companies were failed to define

the capacities of customers, which causes the information asymmetry. Thus, it results in customer expectation exaggeration. (Mimoun, Poncin & Garnier, 2012.) Furthermore, some of the virtual agents behave lack of reciprocity, which means they did not display human embodiment and cannot recognize customer frustration. It makes chatbots always giving customers negative impressions, like cold, untrustworthy, incompetent, etc. (Brave & Nass, 2002). Besides, individuals prefer chatbots to behave more friendly. Users do not like chatbot behaviors with lower positive emotions, fewer assents, and impolite words because they think these behaviors express negative emotions (Skowron et al., 2000). Last but not least, when chatbot answers do not fit with users' questions, users may produce negative feelings of this kind of technology, like "dumb", "impolite", and "rude" (Jenkins et al., 2007). Based on all the reasons which have been listed above, it proved that when chatbots cannot meet customers' expectations, the dissatisfactory online service encounters might be generated (Feine, Morana & Gnewuch, 2019).

The best way to handle different chatbot service encounters is to balance human and technology input, because technology may not always be the best option (Larivière et al., 2017; Frey & Osborne, 2017). The existing research shows that one of the best solutions for human and machine services is their collaboration. Collin's (2018) article mentioned one example of human-machine collaboration. Garry Kasparov was the best chess player between 1986 and 2005, but he was lost a chess game to a computer program from IBM in 1997. Later, he tried to cooperate with machines, and this cooperation shows that when human is working together with machines can beat the singular machine in every chess game. Collin's research suggests that technologies should augment but not replace humans because this type of combination can improve the efficiency of both humans and machines (Jarrahi, 2018; Tripathy, 2018). Feine, Morana and Gnewuch (2019) research stated that customers used to express their frustrations in the text they wrote, and it suggests that chatbots can use sentiment scores to detect users' feelings (identify whether customers with negative emotions or not). In this way, service providers can recognize dissatisfaction moments on time and reduce service failures, such as transferring the conversation to human employees before the service failure happens. Besides, the Uncanny Valley theory suggests that virtual agents or chatbots behaving too humanlike or too unhumanlike will both cause negative results (Groom et al., 2009). It is understandable that some companies might

pursue humanlike chatbot services because they think it can please their consumers. For example, Twitter’s chatbot could not be distinguished from humans by users, and the image for this chatbot is credible, attractive, and efficient (Edwards et al. 2014). However, too much humanlike causes uncomfortable feelings for users (Groom et al.).

One issue for the existing studies is that consumer satisfaction or dissatisfaction results from chatbot service encounters always cannot be retrieved on time because of the rapid development of technologies, and it is difficult to get the newest consumer-chatbot dialog corpus. (Veerhagen et al., 2014.) Thus, the value of this study is highlighted.

### 3.5 Integrative framework: Chatbot service encounters in online customer service experiences

The literature review above primarily focuses on online customer service experiences and chatbots by discussing chatbot service encounters. The author summarized some critical points in Table 1 with four columns. These points consist of some essential concepts, relationships between different concepts, and factors that affect chatbot service encounters from the existing studies.

**Table 1. Summary of the literature review about chatbot service encounters in OCSEs.**

Concepts	Definitions	Related articles	Descriptions
Online customer experience (OCE)	It means “ <i>a holistic and subjective process resulting from interactions between consumers, shopping practices, and the online environment.</i> ”	<ul style="list-style-type: none"> <li>• Trevinal and Stenger (2014, p. 324)</li> <li>• Carbonne and Haeckel (1994, p. 9)</li> </ul>	It formed by people’s encounter with products, services, and businesses which emphasizes online interactions between different players. In this study, the interaction is between consumers and chatbots.
Online customer service experience (OCSE)	It means “ <i>the customers' overall mental perception of their interactions with the online service provider and other customers expressed in</i>	<ul style="list-style-type: none"> <li>• Klaus (2013, p. 448)</li> </ul>	This study focus on the relationship between consumes and online service providers (chatbot).

*its dimension's  
functionality and  
psychological factors".*

Customer satisfaction	It means product/service performance perceived by consumers higher than their expectations.	<ul style="list-style-type: none"> <li>• Oliver (1981, p. 27)</li> <li>• Rose, Hair and Clark (2011, p. 32)</li> </ul>	It is the result of the OCEs.
Service encounter 2.0	It means the "customer-company interaction which interrelated to technologies, human actors, physical/digital environment, and company/customer processes".	<ul style="list-style-type: none"> <li>• Larivière et al. (2017, p. 2)</li> </ul>	It emphasizes interaction between the customer and the company which also related to the technologies.
Online encounter satisfaction	It is the measure of consumers' satisfaction in transactions.	<ul style="list-style-type: none"> <li>• Chan, Barnes and Fukukawa (2016, p. 608)</li> </ul>	Online encounter satisfaction positively affects consumer's overall satisfaction toward a company.
Satisfactory chatbot service encounter	It means consumer satisfaction in service processes under the interactions between chatbots and consumers.	<ul style="list-style-type: none"> <li>• Wolfinbarger and Gilly (2003, p. 196)</li> <li>• Verhagen et al. (2014, p. 539-540)</li> </ul>	<p>Sources of satisfactory chatbot service encounters:</p> <ul style="list-style-type: none"> <li>• Information comprehensiveness.</li> <li>• Service process efficiency.</li> <li>• Chatbot's friendliness.</li> <li>• Chatbot's professionalism.</li> </ul>
Dissatisfactory chatbot service encounter	It means consumer dissatisfaction in service processes under the interactions between chatbots and consumers.	<ul style="list-style-type: none"> <li>• Mimoun, Poncin and Garnier (2012, p. 609-610)</li> <li>• Brave and Nass (2002, p. 54)</li> <li>• Skowron et al. (2000, p. 345)</li> <li>• Jenkins et al. (2007, p. 82)</li> </ul>	<p>Sources of dissatisfactory chatbot service encounters:</p> <ul style="list-style-type: none"> <li>• Lack of online interaction.</li> <li>• Behaves lack of reciprocity.</li> <li>• Impolite expression.</li> <li>• Not able to manage all the information.</li> <li>• Failure to understand customers.</li> <li>• Behaves aggressively.</li> <li>• Uncomfortable interaction.</li> </ul>

		<ul style="list-style-type: none"> <li>• Answers do not fit the question.</li> </ul>
Results of the service encounter satisfaction	<ul style="list-style-type: none"> <li>• Oliver (1997, p. 5)</li> </ul>	Relevant factors: <ul style="list-style-type: none"> <li>• Word-of-mouth.</li> <li>• Customer loyalty.</li> <li>• Repurchase behavior.</li> </ul>

The author picked the sources of different chatbot services from Table 1 and classified them into three dimensions based on their properties. The three dimensions are the properness of reply, intelligence, and the properness of behavior based on commonalities. The properness of reply means to answer questions suitably or correctly (Cambridge dictionary, 2020c). The intelligence means “*the ability to learn, understand, and make judgments or have opinions based on reason*” (Cambridge dictionary, 2020a). The properness of behavior means to behave suitably or correctly (Cambridge dictionary, 2020b). (Table 2)

**Table 2. The sources of different chatbot service encounters and its dimensions.**

Concepts	Sources of different chatbot service encounters:	Dimensions
Satisfactory chatbot service encounter	N/A	1) Properness of reply
	<ul style="list-style-type: none"> <li>• Information comprehensiveness.</li> <li>• Service process efficiency.</li> <li>• Chatbot’s professionalism.</li> </ul>	2) Intelligence
	<ul style="list-style-type: none"> <li>• Chatbot’s friendliness.</li> </ul>	3) Properness of behavior
Dissatisfactory chatbot service encounter	<ul style="list-style-type: none"> <li>• Impolite expression.</li> </ul>	1) Properness of reply
	<ul style="list-style-type: none"> <li>• Answers do not fit the question.</li> <li>• Not able to manage all the information.</li> <li>• Failure to understand customers.</li> </ul>	2) Intelligence
	<ul style="list-style-type: none"> <li>• Lack of online interaction.</li> </ul>	3) Properness of behavior
	<ul style="list-style-type: none"> <li>• Behaves lack of reciprocity.</li> <li>• Behaves aggressively.</li> <li>• Uncomfortable interaction.</li> </ul>	

The previous literature has stated the relationship between different concepts, which are listed in Table 1. First of all, the OCSE is a part of the OCE, which is formed by the interaction between different players (products, services, and businesses). It

emphasizes the overall perception of the interactions between consumers, shopping practices, and the online environment. This study focuses on the interactions between chatbots and consumers. Thus, in Figure 1, the whole frame is the “online customer service experience”, and chatbots & consumers are on the side of the frame. Second, customer satisfaction is the result of OCSEs, and online encounter satisfaction positively affects consumers’ overall satisfaction toward companies. Therefore, consumer satisfaction (overall) is placed as the result of the chatbot service encounter. Third, the service encounter 2.0 emphasized the interaction between the company and its customers, and this study focuses on the chatbot. So, the chatbot service encounter is put in the middle of the figure with a short introduction (interaction process). Furthermore, the author puts the sources of satisfactory and dissatisfactory chatbot service encounters, which are summarized in Table 2 on the left side of the figure and will complement it with the findings from this study in the conclusion chapter. Finally, the service encounter satisfaction has an impact on consumer behaviors. So, the author puts this element as a result of customer satisfaction (overall).



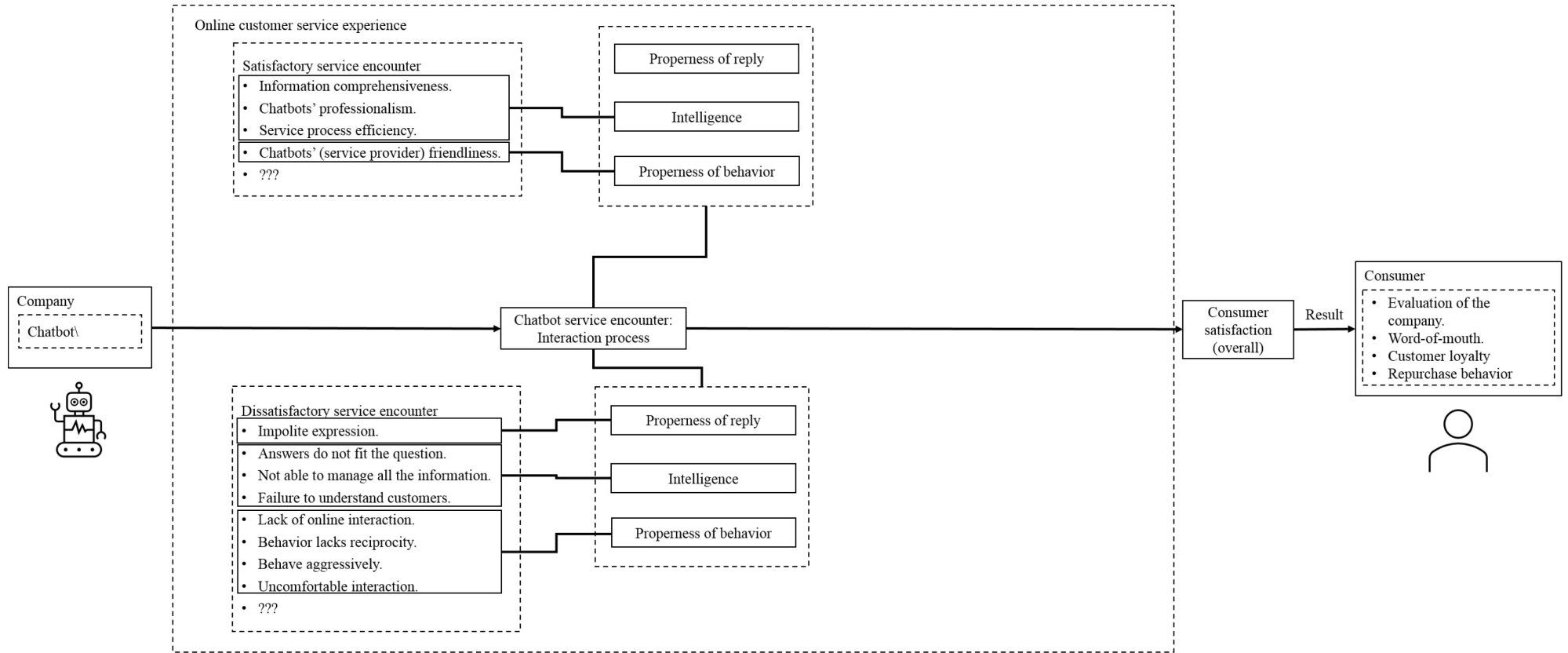


Figure 1. Integrative framework: Chatbot service encounters in OCSEs.

## 4 METHODOLOGY

This chapter presents the research methods employed in this study. The data collection methods are presented at the beginning, which includes CIT, focus groups, and the combination of these two methods. The next section is about the data collection process, which is demonstrated step by step. Then, the last section presents the data analysis method with the data sorting process.

### 4.1 Combination of CIT and focus group methods

There is no doubt that the chatbot service encounter is not a merely “yes” or “no” question. Thus, this study chooses a qualitative research methodology and combines two qualitative data collection methods – CIT and Focus Group. Besides, in order to ensure that the critical incidents are able to meet the requirements of this study, the author used the pre-questionnaire as a supportive method.

The critical incident technique (CIT) can be dated back to 1954; Flanagan (1954) used this method found out the requirements of an activity. Then, this method was proved to be suitable for the research that aims to increase knowledge and to understand phenomena by Bitner, Booms and Tetreault in 1990. Later, Butterfield et al. (2009) stated a similar opinion that “*CIT explores what helps or hinders in a particular experience or activity*”. All in all, CIT is suitable to acquire information about behaviors and experiences, which results in satisfaction or dissatisfaction (Viergever, 2019). Thus, the author applied CIT in this study.

The last paragraph mentions that Bitner, Booms and Tetreault (1990) used CIT to assist them in understanding the service encounters and declared that this is the most appropriate way to understand customer satisfaction and dissatisfaction in service encounters. In their study, the critical incidents were defined as customers’ particularly satisfying and dissatisfying memorable interactions with human employees. They did interviews for their samples, which allows the interviewer to observe responders’ behaviors. Then, they used a content analysis method to analyze the stories collected from interviewees. Different from Bitner, Booms and Tetreault’s study, machines

replaced human employees in this study. Thus, the author redefined the critical incidents in this study as satisfactory and dissatisfactory experiences with chatbot services, based on the aim to find out the theoretical framework of the advanced chatbot service encounters in online customer service experiences. It has a few requirements for the incidents 1) Involving consumer-chatbot interaction, 2) from consumers' perspective, 3) incidents are very satisfying or dissatisfying experiences, 4) the description includes enough details. Besides, this study uses focus groups instead of interviews, and reasons are presented in the next few paragraphs.

The focus group is a suitable way to collect information about individuals' experiences, and it is a way to elicit participants' preferences about one thing (Hines, 2000). It enables researchers to involve in the data collection process to get more insights into the data (Yin, 1994). The advantages of focus groups are: 1) it is a low-cost way to collect rich data, 2) it is flexible, 3) it can stimulate the respondents during the discussion, 4) it able to aids recall, 5) the researcher is able to accumulate responses from all participants, 6) and the experiences can be shared by both groups and individuals. (Hines, 2000.) These benefits can help the author to understand the critical incidents in this study better because of the incidents in this study are generally about consumer experiences. The disadvantages of focus groups are: 1) it requires the moderators able to manage the process, 2) the process might be deteriorated (such as dominate by one individual), 3) it might be difficult to manage sensitive questions, 4) the process might be misleading. These disadvantages reflect the moderator's importance and require them to have a clear mind about what information is necessary for the study because the focus group discussion process affects data's qualities. (Hines.)

For this current study, the author is the moderator for the data collection process. The focus group discussions were organized online due to the coronavirus situation, and participants cannot meet up around a table. Online focus groups are an alternative way when the face-to-face focus group is not available (Tenney, 2016). It means to operate the focus group in a virtual discussion room, participants in the "virtual room" can answer and interact with the moderator and other participants (Hancock, 2017). The online focus group is not a new way under the background of Web 2.0. It can help researchers to save costs, organize in different locations, attract specific participants,

etc. The tasks of moderators for the online focus group are compared to the offline (face-to-face) focus group. However, the online environment might affect nonverbal communications. (Stewart & Shamdasani, 2017.) In order to conquer this disadvantage, the author uses a synchronous type of online focus group with real-time video discussions instead of a simple voice meeting, which allows every participant can see each other.

Except for the moderator, the properness of the sample size is another crucial factor affecting the researchers to collect rich data and avoid redundancy. It has one phrase called "data saturation" or "thematic saturation", which refers to a "special point" during the data collection process. This point happens when the data starts to repeat, which means the rest of the information is not necessary anymore. The saturation consists of "code saturation" and "meaning saturation". For both saturations, the majority codes (deductive codes) and information are generated from the first focus group, which is clearly decreasing in the following groups. After the second group, no new deductive codes are appearing anymore. (Hennink, Kaiser & Weber, 2019.) For this current study, the author decided to use three focus groups due to time and cost limitations (the author had very limited time to collect the data). Another reason is that the three groups are enough to obtain most of the information.

All in all, this study combined the CIT and focus group method to collect more comprehensive data. Participatory research (such as focus group) is seen as a helping hand for CIT, and the critical incident can assist researchers in understanding and guiding the focus group discussion (Getrich et al., 2016). In this study, focus groups are helping the author to have a deep understanding of critical incidents, and critical incidents are assisting the author in comprehending consumer experiences with chatbots better. Besides, in order to ensure the critical incidents can meet the requirement of this study, all the potential participants were required to complete an online pre-questionnaire with one opening question on the Microsoft Word before the focus group interviews. This question requires the participants to write down their satisfactory and dissatisfactory experiences with chatbots (Appendix 1). More details about the data collection process are presented in the next sub-chapter.

## 4.2 Data collection process

The sample of this study consist of 12 participants from the age range between 18-35 years old (no nationality restrictions). The author decided to focus on the younger generation because younger netizens are more skillful to access advanced technologies. Also, the author is more interested in understanding the younger generation's behavior. All of these 12 participants have interactive experiences with a text-based utilitarian chatbot before. The interactions can be every type of service among all industries, such as F&Q chatbots, online shop chatbots, working assistants, etc.

As mentioned in the last section that before the focus group discussion, the author has sent the pre-questionnaire (Appendix 1) to the potential participants to make sure all the participants' experiences able to match the requirements for the critical incidents for this study. It could be seen as the groundwork for engagement of the focus group interview. A total of 12 online pre-questionnaire, with 24 incidents were collected. After collecting all the pre-questionnaire, the 12 participants were divided into three focus groups (the Chinese participants were in the same group because it is easier to conduct the focus group discussion on WeChat), and each of them consists of 4 participants.

For each focus group, the contents of the discussion consisted of 2 parts. The first part focuses on participants' positive experiences with chatbots, and the second part focuses on participants' negative experiences with chatbots. Both of them have the same discussion process that participants share their own experiences with chatbots to others first with the same requirements 1) involving consumer-chatbot interaction, 2) from consumers' perspective, the incidents should be very satisfying or dissatisfying, 3) and the description should cover enough details. Enough details mean the experiences include "what types of industry/product", "why you chat with chatbots", "how was the interaction going", and "what the chatbot did make you feel good/bad". In general, it means to share what they have written on the pre-questionnaire. Then the moderator guided the focus group discussion with some semi-questions (see Appendix 2).

All the focus group discussions were held in English or Chinese (then translate to English), and each of them took about 60 minutes. They were conducted on Zoom and WeChat through video calls and were recorded by the author. Table 3 concluded the information for each focus group discussion.

**Table 3. Information for the focus group discussions.**

Group number	Group size	Conducted date	Conducted channel	Language	Duration
Group 1	4	11.04.2020	Zoom	English	70 minutes
Group 2	4	12.04.2020	WeChat	Chinese	60 minutes
Group 3	4	19.04.2020	Zoom	English	60 minutes

### 4.3 Data analysis method

After the focus group discussions, the author transcribed all screen recordings (focus group discussions) to text files (Microsoft Word), and each document has approximately 6-9 pages. Finally, 22 valid critical incidents and two invalid incidents were collected. Ten of them are favorable experiences with chatbots, and 12 of them are unfavorable experiences with chatbots. Then, the author analyzed the data on NVivo.

The data analysis method used in this study is the thematic data analysis method. Getting the "theme" is the central part of this method with the process of identifying, analyzing, and interpreting the meaning of the data. This process can offer researchers a systematic procedure to generate codes (the smallest units of the theme) and themes. (Clarke & Braun, 2017.)

Concretely speaking, the first step is recurring themes that are most relevant to this study (cleaning the data). The author divided the data into two big themes first (satisfactory and dissatisfactory). Then put the data into finer themes based on the nature of chatbot's behaviors, which caused the satisfactory and dissatisfactory chatbot service encounter, such as consumers ask for product information, consumers ask for cloth suggestions, etc. It should be noted that based on the goal of answering the research questions, some data were sacrificed during this data cleaning process. Still,

the author tried to keep the data as comprehensiveness as possible. The whole theme analysis process was a careful reading process, and the similarities between the different themes appeared gradually. Based on the similarities of different themes, the author sorted the data into different groups step by step.

The data sorting process logic from Bitner, Booms and Tetreault's (1990) study was applied in this current study. The logic of this data sorting process is like a "decision tree", which means using questions to refine the themes step by step until it is not necessary to be refined anymore. In a word, an iterative process that divided incidents into different groups until consensus achieved. As Figure 2 shows, the first question node applied by the data sorting process was "is there a request for chatbot services from consumers". Answering no goes to the branch about the nature of unprompted chatbot actions (Group 3) with three leaves. Answering yes goes to the next question node "is there a consuming behavior happened". This question came out from the data that did not go to Group 3. Answering yes means consumer purchased products/services from companies already and then goes to a branch about the nature of after-sales service (Group 1) with two leaves. Answering no goes to a branch about the nature of needs (Group 2) with three leaves. It is necessary to mention that all the question nodes came from the similarities among different themes. All themes were experienced reading, coding, re-coding, sorting, and re-sorting because the author always found something new during the data sorting process. In the end, the eight categories were defined, which are discussed concretely in the next chapter.

After all groups and categories were generated, the author did a simple quantitative descriptive analysis for the results (Table 4 below) to make the data more descriptive. The table consists of groups and categories' information, proportions of satisfactory and dissatisfactory incidents under each group, and total proportions for each group. It provides a very straightforward view of the incident structure and provides insights into the research questions.

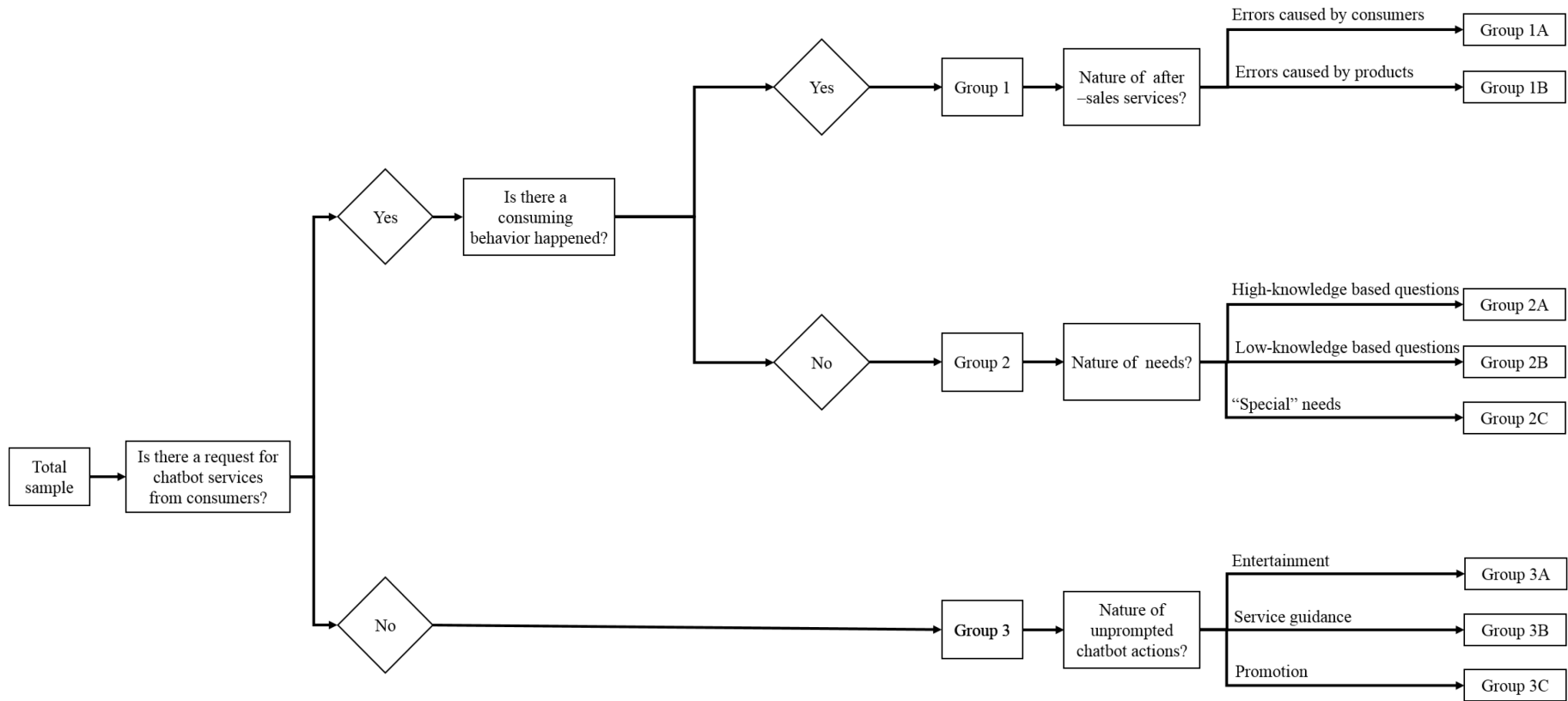


Figure 2. Incident sorting process (in the chatbot situation).



## 5 FINDINGS OF THE EMPIRICAL ANALYSIS

This chapter discusses the findings derived from empirical analysis. The structure of this chapter is associated with research questions. Each sub-chapter corresponds to a sub-question and merge all parts can answer the main research question. The first section generally describes the groups and categories of the data from the focus group discussions into Table 4. It includes the examination for both main groups and categories (under the main groups). The sample incidents within each category are presented too. The second section aims to clarify the findings by illustrating the sources of different types of chatbot service encounters and presents satisfactory and dissatisfactory outcomes separately. The third section generalizes the conclusions of the section of 5.1 and 5.2 to Table 8 and also concluded the sources of different types of chatbot service encounters into three dimensions, which have been introduced to readers in Chapter 3. Based on the findings from the first three sections, one summarizing part of the empirical findings of the sources of different types of chatbot service encounters is presented in the last section of this chapter.

### 5.1 Critical incident classification of chatbot service encounters

This section describes the groups and categories of the data from the focus group discussions. The proportions are shown in Table 4, which aims to provide insights into the data and research questions.

**Table 4. Group and category classification by type of incident outcome.**

Group and category	Type of incident outcome					
	Satisfactory		Dissatisfactory		Row total	
	No.	%	No.	%	No.	%
<b>Group 1: Chatbots response to after-sales services.</b>						
A. Response to admitted consumer errors (caused by consumer).	0	0.0	2	16.7	2	9.1
B. Response to admitted company errors (caused by company/product).	2	20.0	5	41.7	7	31.8
<b>Subtotal, group 1</b>	<b>2</b>	<b>20.0</b>	<b>7</b>	<b>50.4</b>	<b>9</b>	<b>40.9</b>

---

<b>Group 2: Chatbots response to consumers' needs.</b>						
A. Response to inquiries from consumers for some special fields (high knowledge base).	2	20.0	1	8.3	3	13.6
B. Response to the basic questions (low knowledge base).	2	20.0	2	16.7	4	18.2
C. Response to "special" needs from consumers	2	20.0	0	0	2	9.1
<b>Subtotal, group 2</b>	<b>6</b>	<b>60.0</b>	<b>3</b>	<b>25.0</b>	<b>9</b>	<b>40.9</b>
<b>Group 3: Unprompted chatbot actions.</b>						
A. Pop up for entertainment reasons.	1	10.0	1	8.3	2	9.1
B. Pop up for service guidance.	1	10.0	0	0	1	4.5
C. Pop up to promote products.	0	0.0	1	8.3	1	4.5
<b>Subtotal, group 3</b>	<b>2</b>	<b>20.0</b>	<b>2</b>	<b>16.6</b>	<b>4</b>	<b>18.1</b>
<b>Column Total</b>	<b>10</b>	<b>45.5</b>	<b>12</b>	<b>54.5</b>	<b>22</b>	<b>100.0</b>

---

### 5.1.1 Major groups of chatbot service encounters

As Table 4 shows, the data's initial classification resulted in three groups of chatbot behaviors. It covered all satisfactory and dissatisfactory incidents collected from the focus group discussions (the number of percentages is kept one decimal place).

*Group 1. Chatbot responses to service failures.* This group is roughly about the after-sales services, which means the consumer has consumed products or services from the company and then looking for the guarantee, maintenance, and preparation, etc. Good after-sales services are critical to enhancing customer satisfaction and developing a long-term relationship with consumers (Alshare, 2020). This group consists of two types of after-sales services. The first type is about consumers looking for after-sales service due to problems caused by companies or their products/services, such as quality issues. The second type is about consumers looking for after-sales service due to problems were caused by themselves, such as subscribed something by mistake and want to cancel it. Under these situations, consumers are always required to ask either an explicit or inferred request for the after-sales services. Many companies use

chatbots as representatives to solve this kind of problem. Therefore, the responses/replies from chatbots determine customer satisfaction and dissatisfaction.

*Group 2. Chatbot responds to consumer's needs (before consuming products/services from companies).* Consumers are always interacting with a chatbot based on their needs. Thus, chatbots are required to respond to these consumer needs. The content of the chatbot answer determines whether consumers satisfy or dissatisfy for this chatbot service. The incidents in Group 2 are related to the range of chatbot's responses. It includes 1) response to inquiries from consumers for some special fields (high knowledge base) 2) response to the basic questions (low knowledge base), 3) and response to “special” needs from consumers (with some special requirements, such as based on consumer's requirements to provide them some customized recommendations). Whether the contents of replies from chatbot able to answer consumers' questions associated with consumer satisfaction and dissatisfaction.

*Group 3. Unprompted chatbot actions.* From consumers' point of view, these chatbot behaviors are unexpected and unrequested at all. In this study, satisfactory incidents represent the unprompted chatbot services pleased consumers, whereas dissatisfactory incidents mean the unprompted chatbot services cause consumers to generate negative feelings. This group consists of three types of unprompted chatbot actions: pop up for entertainment reasons (to activate the atmosphere), pop up for service guidance, and pop up to promote products. Incidents in this group are not triggered by the core products or consumers' special needs/requests but only triggered by chatbot behavior itself. Thus, chatbot behaviors determine if consumers satisfy or dissatisfy with their experiences.

### 5.1.2 Chatbot service encounter segmentation

There are eight categories in three major groups mentioned above for both satisfactory and dissatisfactory incidents. The last section introduced the frequency of occurrence of these incidents, and this section presents sample incidents within these eight categories (Tables 5, 6, and 7).

*1A. Response to admitted customer errors (caused by consumers).* It means consumers have already purchased or consumed products or services from the company already, and then they look for services from companies due to consumers' subjective behaviors caused problems. For example, a consumer purchased something online and then feel like he/she does not want this product and want to return it, or a consumer subscribed something by mistake and want to cancel it. All participants in this group are eager to solve problems in a short time. Consumers can be very dissatisfied with the situation: 1) chatbots unable to provide a solution instantly for the problem, 2) or consumers originally had a bad impression for this brand/product. Nonetheless, there are no favorable incidents in this category.

*1B. Response to admitted company errors (caused by company/product).* It means consumers have already purchased or consumed products or services from the company, and then they look for services from companies due to the problems caused by products or services. These requests might be related to product quality issues, reclamation issues, etc. Consumers can be very satisfied with the situation: 1) chatbots do not need to provide perfect responses to consumers' questions because consumers agree to fill the questionnaire/form from chatbots and to wait for human responses, 2) or agree with chatbots give and general answers first and contact with human services later. However, consumers can be very dissatisfied with the situation: 1) when chatbots are from big brands with inferior performances, 2) chatbots can understand consumers' questions then provide wrong answers, 3) or chatbots keep providing consumer options to confirm the question, but none of them are relevant to their questions no matter how they change the ways to ask. It indicates consumers' primary impressions for a brand or a company has a potential impact on their attitudes toward their experiences with chatbots. (Table 5)

**Table 5. Group 1: Chatbot response to after-sales services.**

	Incident	
	Satisfactory	Dissatisfactory
<b>A. Response to admitted consumer errors (caused by consumer).</b>		
	N/A	I was looking for the customer services for a chat tool because I scanned a wrong QR code

---

by mistake and want to check if it will bring any risk for my account. I interact with their chatbot, and the whole process was lengthy. I entered the chatbot interface, again and again, the chatbot greeting with me again and again. I described my question in different ways, and the chatbot not able to identify my problem and only provide me urelement options and answers. Then I ask the bot for human services, and the bot offers me a questionnaire to fill, I filled it and sent. Then it reminds me the process will take sometimes. I was furious about that because I think it is a big company and should have good services. (Focus group 2, 6<sup>th</sup> participant, 12.04.2020)

**B. Response to admitted company errors (caused by company/product).**

I wanted to make a reclamation for my flight ticket. It has a chatbot on that company's website, I ask the chatbot for help, and the interaction went very smoothly. The chatbot guided me through asking me questions, then I would answer, and then it would automatically give me another question to process my request. I was surprised by how good it worked. Only a few minutes, I got the link to do the reclamation. (Focus group 3, 12<sup>th</sup> participant, 19.04.2020)

I received an unknown bill from the bank I used to use, and I want to check why I have that bill. Then I asked their online chatbot, which tied with my bank card. However, the chatbot replied to me to contact the human services, or I can choose the further way for another option. I tried, but it is useless. I am abroad, so it is difficult to call them. I was angry because I only want a simple bill history, and they are the biggest bank in my hometown. (Focus group 3, 9<sup>th</sup> participant, 19.04.2020)

---

*2A. Response to inquiries from consumers for some particular fields (high knowledge base).* For this kind of request, chatbots are always expected to have some professional knowledge for one specific area, such as questions about finance, technical issues, tax issues, etc. It indicates that this category is more focused on knowledge transformation. Consumers prefer to evaluate incidents in this category holistically, such as the professional suggestions from chatbots about financial products, which helped consumers made money generates satisfactory consumer experiences. The content of

answers with a high-knowledge base perceived by consumers from chatbots determined whether consumers were satisfied. For example, consumers want some professional knowledge and suggestions to help them choose the proper financial products; employees have some technical problems and need professional solutions, etc. Consumers may remember the encounter as very satisfactory if 1) responses from chatbots are evident with both figures and text introductions; 2) or chatbots are not able to answer consumers' questions directly, but they provide some potential options, and one of them is relevant to the question is enough to generate a favorable incident with satisfactory sensations. In contrast, consumers may remember the encounter as very dissatisfactory if 1) failure to identify questions, 2) or provide unnecessary answers that caused time-wasting.

*2B. Response to the basic questions (low knowledge base).* This category is about answering consumers' simple questions that do not need professional or scientific replies. The ways chatbots reply to consumers' questions are able to determine consumer satisfaction. Providing simple answers is enough to create favorable service encounters. If chatbots respond in a friendly way, consumers' positive attitudes might be enhanced. In contrast, the unfavorable service encounter will be generated if chatbots are not able to identify the keywords from consumers no matter how they changed ways to ask and only offering them the same options.

*2C. Response to "special" needs from consumers.* This category requires chatbot to answer questions based on consumers' particular information, which is more like "customized" chatbot services. The "special" needs are not necessarily focused on the core product, such as a chatbot from an online clothing shop that provides consumers with suggestions about their clothing style based on their preferences. Satisfactory chatbot service encounters are generated when: 1) the recommendations from chatbots able to match with consumers' questions, 2) if chatbots not able to handle the particular need, then transfer consumers' requests to human services. Interestingly, there is no dissatisfactory incident for this category. (Table 6)

**Table 6. Group 2: Chatbot response to consumers' needs.**

Incident	
Satisfactory	Dissatisfactory
<b>A. Response to inquiries from consumers for some special fields (high knowledge base).</b>	
<p>I have questions for my work which require some professional knowledge. I typed my keywords, and the chatbot replied to me 20 relevant questions instant. I found the question I want to ask easily with multiple solutions. The answer was very clear with both picture and text guidance. I can easily understand as a freshman. (Focus group 2, 7<sup>th</sup> participant, 12.04.2020)</p>	<p>Due to the impacts of the coronavirus, the tax rate was adjusted, I want to consult a straightforward question about the start and end dates. The chatbot only rigidly replied to me with a few options he (chatbot) thought are relevant to my inquiry. I tried to ask differently, but it always provides me the same answer. (Focus group 1, 3<sup>rd</sup> participant, 11.04.2020)</p>
<b>B. Response to the basic questions (low knowledge base).</b>	
<p>I played an online video game before, and it has a chatbot assistant in the game. I was looking for one specific non-player character (NPC) in the game. To save time, I asked the chatbot for information by type a few words, and then she provided me an answer about where I can find that NPC in a charming way. (Focus group 1, 2<sup>nd</sup> participant, 11.04.2020)</p>	<p>I was looking for a specific piece of clothing on a brand's website. I interacted with the chatbot on their website for information. However, the chatbot seems to do not understand my question and keeps answering, "Do you mean.....?". And the options provide by the bot never be what my initial question. (Focus group 3, 12<sup>th</sup> participant, 19.04.2020)</p>
<b>C. Response to "special" needs from consumers</b>	
<p>I bought a pant online to save time, and I asked chatbot from that brand to recommend some options for pants based on my preference and size information. The chatbot asked me for the basic information and then provided me some suggestions. The whole process was about 10 minutes in total (plus the checking time). (Focus group 3, 9<sup>th</sup> participant, 19.04.2020)</p>	N/A

*3A. Pop up for entertainment reasons.* Incidents in this category reflect consumers' attitudes towards the situation when chatbots are jumping out without asking with the aim to interact with consumers for entertainment instead of asking questions. Interestingly, participants from different groups mentioned the same chatbot from the same chat tool with different attitudes. Satisfactory encounters are associated with consumers feel pleasant due to chatbots brought them fun. Dissatisfactory encounters are associated with consumers feel this kind of chatbot behaviors are impolite (jump out without asking). However, consumers' attitudes toward this kind of chatbot behavior might be affected by situations.

*3B. Pop up for service guidance.* Chatbots in this category are working together with human services. Consumers guided by chatbots to finish the simple steps to help both themselves and human employees save time. This category emerged only for satisfactory encounters. However, it is challenging for the author to conclude that consumers prefer this kind of chatbot service, as this study has a limited sample size.

*3C. Pop up to promote products.* Contrary to the previous group, this category only includes dissatisfactory incidents. The dissatisfactory encounters result when chatbots are popping up to promote products without asking. This kind of chatbot behavior might change consumers' satisfactory experiences to dissatisfactory experiences. However, it is similar to the last category (3B) that it is difficult to conclude because the sample size is small. (Table 7)

**Table 7. Group 3: Unprompted chatbot actions.**

Incident	
Satisfactory	Dissatisfactory
<b>A. Pop up for entertainment reasons.</b>	
It has a chatbot on a chat tool, the chatbot asked us to play a game, and I think the chatbot is very intelligent. My friends and I were in a group chat, and the chatbot brought joy for us. It increased the fun of our group chat. (Focus group 1, 1 <sup>st</sup> participant, 11.04.2020)	I was chatting with my friends in the group, the chatbot suddenly jumped out and asked us to play a game. I was angered because I did not ask the chatbot for a game. Besides, the chatbot used a very formal way to ask us to play a game, and it made me feel



uncomfortable. (Focus group 2, 8<sup>th</sup> participant, 12.04.2020)

**B. Pop up for service guidance.**

I went to a bank to deal with something about my bank account. I went to the number machine to take a waiting number. Suddenly a chatbot popped up. The chatbot greeted me and asked what I want to do. I typed the keywords, and the chatbot asked me to fill a form that helped me solved my problem. I think it helped me to save time. (Focus group 2, 5<sup>th</sup> participant, 12.04.2020)

**C. Pop up to promote products.**

N/A

My experience is with an airline company. I wanted to book a flight ticket, and one chatbot popped up during my booking process and ask me if I want to have something (it is a promotion). Of course, I don't want, I closed the window, but it happened in the next few steps again. It was so annoying, and I just wanted to do it quickly. (Focus group 3, 10<sup>th</sup> participant, 19.04.2020)

## 5.2 Sources of satisfactory and dissatisfactory chatbot service encounters

The three groups and eight categories mentioned above capture the types of chatbot behaviors that lead to very satisfactory and dissatisfactory chatbot service encounters. This section aims to make the results clearer by organizing the fragment source of satisfactory and dissatisfactory chatbot service encounters from the last chapter into different dimensions.

### 5.2.1 Sources of satisfactory chatbot service encounters

The data in Table 4 offers insights into the structure of chatbot service encounters. 20.0% of the satisfactory chatbot service encounters are classified in Group 1, which represents all the incidents relevant to after-sales services. The data reflects that during the interaction process when 1) a chatbot automatically provide another question to precise consumers' requests after consumers typed the keywords (with a clear logic), 2) or if chatbots are not able to provide comprehensive answers, then answering part of questions and offering some potential ideas about what consumers could have a try able to create satisfactory chatbot service encounters. Besides, it is interesting that there are no satisfactory incidents about chatbot response to admitted consumer errors. It means it is difficult for chatbots to handle the situation when errors are caused by consumers, such as subscribed unwanted services, scanning the QR-code by mistake, etc. From a management perspective, it suggests that companies should use human employees to handle the situation when the problem for after-sales services are caused by consumers instead of products/services.

*“I asked chatbot by keywords – a message board for the public account. The chatbot replied to me the message function was maintaining that time, and all the accounts registered around that time not able to use that function, so that function is not available at that time. This reply answered half of my question because it didn't mention precisely since what time. However, the chatbot provides me the way about how to set the message board function. I think the chatbot has clear logic and guide me to the next question.”* (Focus group 2, 6th participant, 12.04.2020)

*“The interaction went very smoothly. The chatbot was asking me questions, then I would answer, and then it would automatically jump to another question to precise my request, and so on. I was actually greatly surprised by how good and how this chatbot worked.”* (Focus group 3, 12th participant, 19.04.2020)

Over half satisfactory incidents are from Group 2 (60.0%), which related to the way chatbots respond to different types of consumer needs. The data reflects that when chatbots 1) provide answers for technical questions, an integrated answer with texts, figures, reference times, etc. together, 2) response to high-knowledge base questions with proper explanations, 3) provide consumers simple answers for their basic questions (low-knowledge base) immediately and correctly, 4) or able to match

consumer needs for customized service requests are able to make consumers remember these kinds of experiences as very satisfying.

*“I asked Wanxiang (the name of the chatbot) for a technical problem by keywords, which is about my user’s ID. The chatbot gave me about ten different relevant question options. Options include the number of how many times this question be asked before. I got the answer to my question easily, and the response consists of both figure and text guidance. It means you can understand it easily even you are a freshman. I think the answer was very clear and understandable.”* (Focus group 2, 5<sup>th</sup> participant, 12.04.2020)

*“I want to buy pants from an online shop, and it has a chatbot assistant asked for my information, such as weight, height, etc. Then the chatbot gave me some recommendations and suggestions for the clothing style. I like its recommendations, and the whole purchasing process only took me 10 minutes.”* (Focus group 3, 9<sup>th</sup> participant, 19.04.2020)

Finally, observation from Table 4 shows that the unprompted chatbot actions contribute 20.0% of satisfactory chatbot service encounters, which means chatbot pop up for service, promotion, or other services without consumer’s asking. The data reflects that when chatbots 1) pop up for guidance purposes, 2) or jump out to increase the lively atmosphere can always bring consumers with satisfactory experiences. However, it has an opposite attitude from participants toward the chatbot pop up to active the atmosphere towards the same chatbot, which indicates consumers’ characteristic affects their attitudes towards the chatbot services. Besides, there are no satisfactory incidents for unprompted chatbot actions with the purpose of promotion, and it suggests companies should not use chatbots to do promotions.

*“I went to a bank, and I have to wait in the queue. It has a service machine with a chatbot in the reception place. The chatbot greeting with me first and I typed to the chatbot what I want to do, then the chatbot replied to me and asked me to fill a form, which helped me solve my problem without the human services.”* (Focus group 2, 5<sup>th</sup> participant, 12.04.2020)

*“In a chat with a group, we found that there was a robot called Xiaobing, and she was a chatbot who can automatically interact with us. She jumped out in the group chat and asked us if we want to play a game with her. We send her an interesting message and wait for her response. To our great surprise, she put amazingly interesting words, saying that we can do a game together.”* (Focus group 1, 1<sup>st</sup> participant, 11.04.2020)

### 5.2.2 Sources of dissatisfactory chatbot service encounters

The classification system also informed the primary source of dissatisfactory chatbot service encounters. The examination of Table 4 reveals that the largest proportion of dissatisfactory chatbot service encounters (58.4%) are linked with Group 1. These incidents reflect which kind of actions from chatbot for the after-sales (after-consume) services cause consumer dissatisfactory experiences. The dissatisfactory chatbot service encounters include 1) the chatbot service interface mistakes cause consumers to enter to the interface again and again, 2) chatbots responses to different keywords from consumers with the same irrelevant answers repeatedly no matter how consumers change the way to ask, 3) chatbots provide consumers wrong information, such as the wrong number to contact the human services, 4) chatbots make too many rounds sub-questions to precise consumers' question, and then provide a very lengthy answer, 5) chatbots did not update its database for the newest product on time caused chatbots not able to answer the questions, 6) or chatbots reply consumers in impolite ways with useless answers. Besides, the last section also pointed out that consumers might be more sensitive toward the negative chatbot service encounters from the big brands, as some participants mentioned in the focus group discussion that "as a big company, they should have capabilities to optimize their chatbot services". The classification system also informed the primary source of dissatisfactory chatbot service encounters. The examination of Table 4 reveals that the largest proportion of dissatisfactory chatbot service encounters (58.4%) are linked with Group 1. These incidents reflect which kind of actions from chatbot for the after-sales (after-consume) services cause consumer dissatisfactory experiences. The dissatisfactory chatbot service encounters include 1) the chatbot service interface mistakes cause consumers to enter to the interface again and again, 2) chatbots responses to different keywords from consumers with the same irrelevant answers repeatedly no matter how consumers change the way to ask, 3) chatbots provide consumers wrong information, such as the wrong number to contact the human services, 4) chatbots make too many rounds sub-questions to precise consumers' question, and then provide a very lengthy answer, 5) chatbots did not update its database for the newest product on time caused chatbots not able to answer the questions, 6) or chatbots reply consumers in impolite ways with useless answers. Besides, the last section also pointed out that consumers might be more sensitive toward the negative chatbot service encounters from the big brands, as some

participants mentioned in the focus group discussion that “as a big company, they should have capabilities to optimize their chatbot services”.

*“I entered the chatbot service interface again and again, the chatbot greeting with me again and again. I described my question, and the chatbot not able to identify what I am asking. Then it provided me some potential options, but none of them is related to my question. This process repeated a few times, and I gave up in the end. Then, I tried to contact human service through the chatbot. The bot did not give me the human service option directly and only provided me a questionnaire to write down my question.”* (Focus group 2, 6<sup>th</sup> participant, 12.04.2020)

*“My new laptops’ camera was not working, and I contacted their customer services. It was a chatbot. The chatbot asked me to type my computer model. After I send the model to the chatbot, it replied me they do not have this model in the system. I was angry. But I think I can understand it because I bought the newest model. Then I look for human services, and they also did not solve my problem. After two months, I asked the chatbot again, and it still not able to solve my problem. I will never buy this brand’s laptop again”* (Focus group 2, 7<sup>th</sup> participant, 12.04.2020)

*“Once, the chatbot sent me a bill, and I do not know where it from. So, I want to check who charged me money. I messaged the chatbot for information, but the chatbot rudely replied to me and asked me to contact the human services. Or I can choose the further way for another option. I tried, but it made me feel like it is useless.”* (Focus group 3, 9<sup>th</sup> participant, 19.04.2020)

In Group 2, Table 4 shows that it has 25.0% of the dissatisfactory incidents are linked with the way chatbots respond to different consumer needs. The dissatisfactory chatbot service encounters will be generated if chatbots are able to identify consumers’ keywords and then provide them some options, but none of them is related to the original questions. This situation works for both low- and high-knowledge base questions. However, there are no dissatisfactory incidents for consumers’ “special” needs, which indicate that it might be easier to please consumers by the customized chatbot services.

*“I only had to write my problem, which the chatbot did not seem to understand. It kept answering that “Do you mean ...?”, of course, the sentence it would propose to me would never be what my initial question has initially meant.”* (Focus group 3, 12<sup>th</sup> participant, 19.04.2020)

*“Every time I ask chatbot questions, it will ask me back with many other questions and request me to choose one of them. Otherwise, the chatbot will not reply to me. The situation always like questions from the chatbots do not match the questions that I want to ask.”* (Focus group 1, 3<sup>rd</sup> participant, 11.04.2020)

Finally, Table 4 reveals that Group 3 has the lowest proportion (16.6%) of dissatisfactory chatbot service encounters, which are relevant to consumers' negative reactions toward unprompted chatbot behaviors. The dissatisfactory encounter includes chatbots 1) pop up to active the atmosphere with very formal words, 2) or chatbots pop up to promote companies' products. It is necessary to note that this kind of chatbot behavior might change the original attitudes from consumers to companies' core products (change positive attitudes to negative attitudes). Besides, there is no dissatisfactory incident related to chatbots pop up with a service guidance purpose, which suggests companies use chatbots to simplify the service process.

*“This chatbot jumped out when I was chatting with my friend in the group without asking, and asked us to play a game. I was angry about that because I did not ask the chatbot. Besides, I feel like this chatbot acted like a machine too much.”* (Focus group 2, 8<sup>th</sup> participant, 12.04.2020)

*“I want to book a flight ticket, and then one annoying chatbot popped up and asked me if I want to have something or want to buy something. Of course, I do not want to buy and closed it. However, it happened in different steps, and I always have to click go back. I was satisfied with their products initially, but I got annoyed during the process, and I really want to do it quickly.”* (Focus group 3, 10<sup>th</sup> participant, 19.04.2020)

### **5.3 Classifying the dimensions of chatbot service encounters**

The purpose of this section is to examine whether the three dimensions about the sources of satisfactory and dissatisfactory chatbot service encounters summarized in the theoretical framework chapter can be used as “generic dimensions”, which can be applied across all industries. The “generic dimensions” can reveal the essence of the sources of chatbot service encounters.

Based on the findings from section 5.2.1 and 5.2.2 about the source of different types of chatbot service encounters (marked by the sub-numbering), the author concludes all of them in Table 8 (columns 1 & 2). Then, through the author’s carefully reading,

sorting, and re-sorting, these sources are concluded into the three dimensions (properness of reply, intelligence, and properness of behavior) found from the theoretical framework chapter based on their similarities.

**Table 8. Classifying the dimensions of chatbot service encounter.**

Categories	Items	Types
Source of satisfactory chatbot service encounter	N/A	1) A proper reply
	<ul style="list-style-type: none"> <li>• For the low-knowledge base question, it provides a correct and concise answer.</li> </ul>	2) Intelligence enough
	<ul style="list-style-type: none"> <li>• If not able to answer the whole question, then answering part of the question and provide some potential idea about what consumers could have a try.</li> </ul>	
	<ul style="list-style-type: none"> <li>• For the high-knowledge base question, it provides a comprehensive answer (figure, text, reference time, etc.).</li> </ul>	
	<ul style="list-style-type: none"> <li>• Based on the data from consumers, it provides a customized answer.</li> </ul>	
	<ul style="list-style-type: none"> <li>• If not able to identify the keyword, then provide further options to precise the keywords in a clear logic.</li> </ul>	
Source of dissatisfactory chatbot service encounter	<ul style="list-style-type: none"> <li>• Pop up to provide guidance.</li> </ul>	3) A proper behavior
	<ul style="list-style-type: none"> <li>• Pop up to activate the service atmosphere in a proper way.</li> </ul>	
	<ul style="list-style-type: none"> <li>• For a normal question, the chatbot provides too many rounds of sub-question with a lengthy answer.</li> </ul>	1) An improper reply
	<ul style="list-style-type: none"> <li>• Reply in an impolite way with a useless solution.</li> </ul>	
	<ul style="list-style-type: none"> <li>• After identifying the keyword, then providing irrelevant options.</li> </ul>	2) Lack of intelligence
	<ul style="list-style-type: none"> <li>• Provide the consumer with the wrong information.</li> <li>• Not able to identify the keyword, then repeat the same answer no matter how consumers change their questions.</li> <li>• Chatbots did not update their database on time, caused not able to answer consumers' questions.</li> <li>• The interface mistakes caused problems.</li> </ul>	

- Pop up to promote products.
  - Pop up to activate the service atmosphere in an improper way.
- 3) An improper behavior
- 

## 5.4 Summary of the empirical findings

This section recapitulates the empirical findings of the sources of chatbot service encounters, which are presented in the previous sections. Also, it compares these findings with the earlier studies, both similarities and differences are found. Besides, some other findings from the empirical study, which are not directly relevant to the sources of different types of chatbot service encounters are presented at the end of this sub-chapter.

### 5.4.1 Summary of the satisfactory service encounter

This study finds that the comprehensive answers for the high-knowledge base questions can lead to consumer satisfaction, which is supporting the ideas from Wolfinbarger and Gilly (2003) and Verhagen et al. (2014) that information comprehensiveness and service providers' professionalism affects satisfactory encounters. This point belongs to the intelligence dimension. Moreover, this study finds that further options to precise the keywords can generate consumer satisfaction, supporting the idea from Wolfinbarger and Gilly (2003) that the efficiency of the service process impacts consumer satisfaction. This point belongs to the intelligence dimension too.

However, it is interesting that no participant considers their satisfactory service encounters are generated from chatbots' friendly, cutie, or fancy answers, which belongs to the dimension of the properness of the behavior. Some participants mention that they do not need chatbots to use cutie ways to tell them that "I cannot solve your problem." It means there is no point in supporting the idea from Verhagen et al. (2014) about how service providers' friendliness affects customer satisfaction. The potential reason for this could be that consumers treat chatbots as machines because participants mentioned in the focus group discussion that "*they do not want to make friends or keep*



*relationships with a machine. Otherwise, others might think they have some mental issue*” (Focus group 2, 12.04.2020). Also, this result related to Mou and Xu’s (2017) study is about how people act differently when they communicate with humans and AI. Besides, this point relevant to Hobbs and Evans’s (1980) studies too, their study pointed out that one of the goal-driven in human-human communication is the relationship goals. The result of the consumer treats chatbots as machines indicated that the goal-driven in HMC is different from human-human communication.

This study found some new sources apart from the existing studies of satisfactory chatbot service encounters based on the critical incidents collected by this study. It includes chatbots 1) provide correct and concise answers for low-knowledge questions, 2) provide one or few potential ideas about what consumers could have a try when it not able to answer the questions, 3) provide customized answers, 4) pop up to provide guidance, 5) and pop up to active the service atmosphere in a proper way are all able to create satisfactory chatbot service encounters. The first 3 points are belonging to the intelligence dimension, and the last 2 points are belonging to the properness of the behavior dimension.

Besides, this empirical study did not find any sources of dissatisfactory chatbot service encounters from the category 2C (response to “special” needs from consumers) and the category 3B (pop up for service guidance). Thus, companies could train their chatbots to handle the recommendation types of questions, especially for online shopping. Due to the participants of this study mentioned that chatbots could, based on their preferences and personal information, provide clothing style quickly. The case of 3B suggests companies can use chatbots to offer service guidance.

#### 5.4.2 Summary of the dissatisfactory service encounter

First, this study finds out that impolite reply with useless solutions causes dissatisfactory service encounters, which assists the idea from Mimoun, Poncin and Garnier’s (2012) opinion about the impolite expression affects consumer attitudes towards their experiences. This point fits into the dimension of the properness of reply. Secondly, the aspects of providing irrelevant options to precise the question and not able to identify the keywords correspond to Mimoun, Poncin and Garnier’s (2012) idea

about fail to understand customers cause dissatisfactory chatbot service encounter. It also fits the idea from Jenkins et al. (2007) that answers do not fit the question lead to negative service encounters. Third, the point about providing wrong information corresponds to Mimoun, Poncin and Garnier's (2012) result that not being able to manage all the information leads to dissatisfaction with chatbot service encounters. The second and third points are belonging to the intelligence dimensions. Fourth, the opinion about pop up to active the service atmosphere in an improper way in line with the view from Mimoun, Poncin and Garnier's (2012) that behaving lack of reciprocity and uncomfortable interaction results in the unfavorable service encounter. It also proves that some people prefer real human services is because human service can identify subtle linguistic cues (Feine, Moorana & Gnewuch, 2019). All in all, the above ideas also echo to Zhu et al.'s (2018) study about the challenges companies face to build their chatbots. Their study stated that lack of dialogue data and poor performance for multi-turn conversations are two main challenges for companies, explaining why chatbots have poor performance.

Nevertheless, there are no participants mentioned in the focus group discussion that the feeling of dissatisfaction is because they lacked interaction with chatbots or the chatbot behave aggressively (both of them belong to the properness of behavior dimension). Sandry's (2018) study is able to explain part of this phenomenon that HMC using more natural/human language than before.

Apart from the existing sources, this study figures out a few new sources for dissatisfactory service encounters. It includes 1) chatbot uses many rounds of sub-question, then provides a lengthy answer (belongs to the intelligence category) 2) and pop up to promote products (belongs to the properness of behavior category).

Furthermore, this empirical study did not find any sources of satisfactory chatbot service encounters from category 1A (response to admitted customer errors) and category 3C (pop up to promote products). The case for 1A implies that companies better to use human employees to handle the questions from the after-sales services caused by consumers. Participants mentioned in the focus group discussion that when they face this kind of problem, they are always in a hurry and want to solve it at once. More generally speaking of the Group 1 (chatbot response to after-sales services) that

over half dissatisfactory chatbot service incidents belong to this group. It means chatbots are not able to deal with this kind of situation well now, and it is better to involve human employees. However, the participants mentioned that in some situations, they are not able to contact the human employee, and chatbots are not able to help them at all. The case for 3C indicates that for companies, it is better not to use chatbot as a promotion tool, as one participant mentioned that he had a good service experience when he was booking a flight ticket. The popped up chatbot changed his attitude from satisfaction to dissatisfaction. Another point from the focus group discussion is about chatbot pop up to active the atmosphere. Companies should optimize their chatbot behavior to fit the situation. Otherwise, it will bring negative feelings to consumers.

In addition, this empirical study also discovered some other interesting findings. Firstly, consumers have a higher expectation for chatbots from a big company/famous brand. A few participants in the focus group discussion mentioned that *“I think it is a big company, so they should have good chatbot service. I am very disappointed with it”* (Focus group 2, 12.04.2020; Focus group 3, 19.04.2020). It reminds the big company to increment their chatbot function. Secondly, consumers still need time to balance technology and their cognition. During the focus group discussion, some participants mentioned that they still prefer human services, and some participants think that chatbots performed better. This finding is in conflict with the previous study that chatbots could be seen as the proper candidate for alternative traditional customer services, especially when the machine performs better than human labor (Brandtzaeg & Folstad, 2017; Huang & Rust, 2018). It indicates that companies should combine these two service methods to please more consumers. Furthermore, through this empirical study, the author argues that the majority of consumers' needs for chatbot services are low-level needs. Due to participants for this study were mentioned that they just want chatbots to answer questions instead of performing like a human (high-level needs). It supports the idea of Maslow's Hierarchy of Needs theory (Maslow, 1987), he stated that people progress on higher-level needs after the lower-level needs are satisfied. As the participants mentioned in the focus group discussion that they do not need a "machine" to use a cute way to tell them that "I do not understand your question". Last but not least, consumers treat machines differently with human employees. The participants mentioned in the focus group discussion that *“I do not*

*want to make friends or keep relationships with a machine. Otherwise, others might think they have some mental issue”* (Focus group 2, 12.04.2020). This result is related to Mou and Xu’s (2017) study that people act differently when communicating with humans and AI. It indicates that the goal-driven in HMC is different from human-human communication due to the finding in Hobbs and Evans’s (1980) study, which has stated that one of the goals of human-human communication is the relationship goal.

## 6 CONCLUSIONS

This chapter answers the research questions at the beginning, and each question is discussed separately. Then, based on the answers for the research questions and some interesting ideas from the empirical study (presented at the end of Chapter 5), the theoretical contribution and managerial implications are demonstrated. The next section illustrates the validity and reliability of this study, which is followed by the limitations of this study. The last section presents suggestions for future studies.

### 6.1 Answers to the research questions

This study aims to discover the theoretical framework of the advanced chatbot service encounters in online customer service experiences by understanding consumers' satisfactory and dissatisfactory experiences with chatbots. The author generated an integrated framework about chatbot service encounters in OCSEs in accordance with both theoretical and empirical studies. Furthermore, this study developed an incident sorting process model for chatbot service encounters. Based on this model, 16 sources of different types of chatbot service encounters were found. This model could apply to future studies or companies to understand the sources of satisfactory and dissatisfactory chatbot service encounters. Also, three dimensions for the chatbot service encounters were found in this study.

The main research question is: *“What is the theoretical framework of the chatbot service encounters in online customer service experiences?”*. The author summarized the existing studies relevant to chatbot service encounters in Table 1 and Table 2. Then, the information presented in these two tables were illustrated in Figure 1 (Integrated framework: Chatbot service encounters in OCSEs). The empirical study found some new sources, which has been concluded in Table 8 with seven sources of satisfactory chatbot service encounters and seven sources of dissatisfactory chatbot service encounters. The differences and similarities about the sources of different types of chatbot service encounters have discussed at the end of the last chapter. In this section, these sources are fitted on the right side of Figure 3 below.

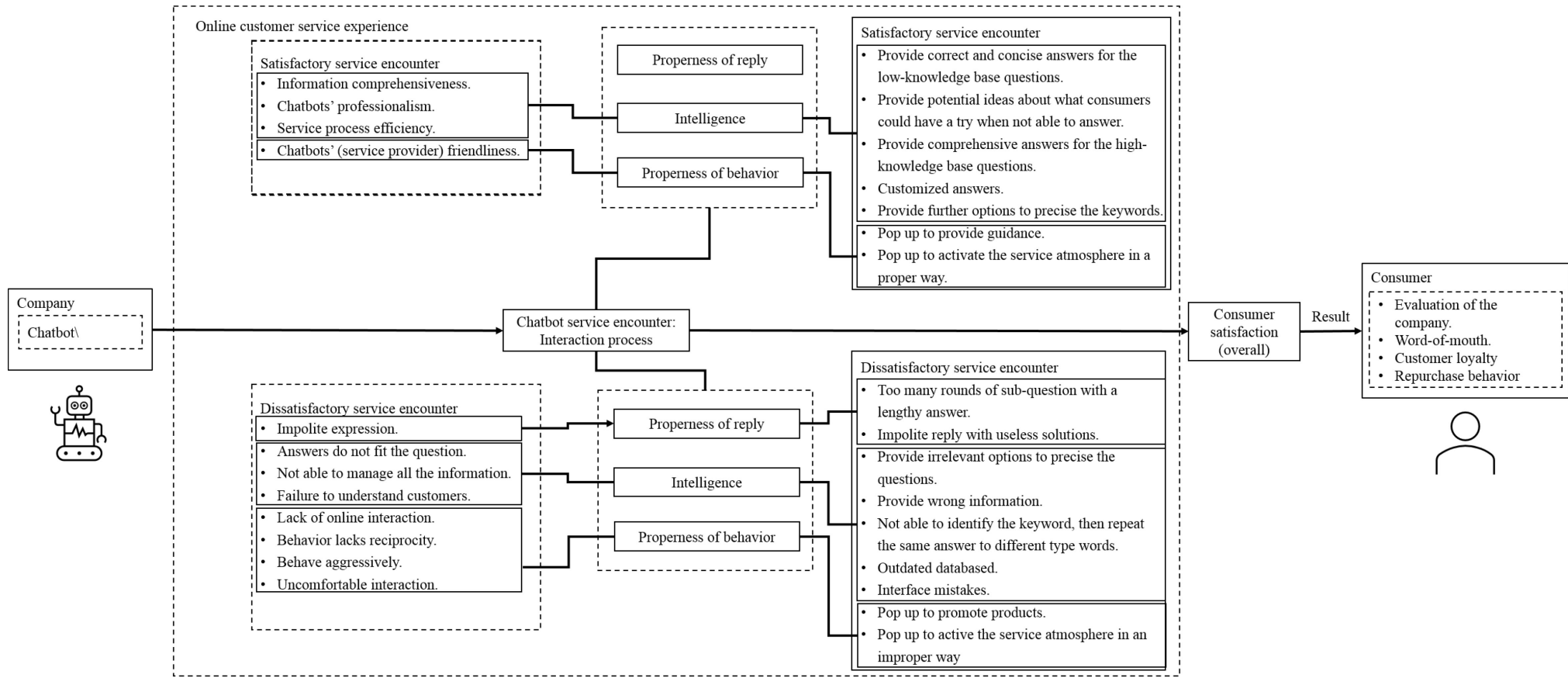


Figure 3. Summary: Chatbot service encounters in OCSEs.

Figure 3 clarifies and provides an overall overview of the “chatbot service encounters in OCSEs”, and it can offer future studies a quick check about the outline of chatbot service encounters. Generally speaking, this framework indicates the relationship between different concepts, which includes chatbots, chatbot service encounters, consumer satisfaction, and consumers. Specifically, the interactions between consumers and chatbots are chatbot service encounters, which are influenced by different factors (such as if the chatbot is intelligent enough). These chatbot service encounters determine consumers’ attitudes and behaviors. Besides, the sources of chatbot service encounters that have been presented in this framework can provide companies with managerial implications when considering developing their chatbot services. The implications could be: focus on what they can do to avoid the dissatisfactory chatbot service encounters, what they can do to enhance their chatbot services, and why chatbot service encounter is essential among others.

After answering the main research question, the three sub-questions are discussed next. It should be mentioned that the answers to these three questions were all demonstrated in the previous section. Thus, the author only discusses them in a condense manner in this section. The first sub-question is “*What is the incident sorting process for the chatbot service encounter?*”. The answer to this question was generated during the data analysis process, and the section of 4.3 has provided the data sorting process in a more detailed way.

Chapter 4 has mentioned that this incident sorting process was inspired by the data sorting logic from Bitner, Booms and Tetreault’s (1990) study. The author developed Figure 2 (incident sorting process for the chatbot services) based on their logic. As Figure 2 shows, this incident sorting process divided the chatbot service incidents into three groups by two questions, which were developed based on the similarities of the data from the focus group discussion. These questions are about: the initiative to make requests and whether the consuming behavior occurred. Groups divided by these questions consist of a few sub-groups used to refine the data to understand these incidents better.

This data sorting process can be used in future studies to classify the chatbot service encounter, or it can be sued by companies to understand and manage their chatbot

services better. For example, through this data sorting process, companies are able to know what types of chatbot services they are providing. They can then check the sources of that kind of chatbot service encounters to get a better understanding of their chatbot services. However, this data sorting process might only work for the utilitarian chatbot services.

The second sub-question is: “*What are the sources of satisfactory and dissatisfactory chatbot service encounters?*”. The answer to this question has been demonstrated in the section of 5.2, and the sources of chatbot service encounters were marked by sub-numbering in the text, like “1)”, “2)”, etc. Then, the author sorted them into Table 8 to have a clear review. As Table 8 shows, the satisfactory chatbot service encounters consist of seven sources, and the dissatisfactory chatbot service encounters comprised of nine sources. These sources suggested that consumers have different expectations for different types of services. For example, consumers prefer a concise answer for a low-knowledge base question but prefer a comprehensive answer for a high-knowledge base question. This situation suggests companies should understand their business type and their consumers’ expectations.

The last sub-question is: “*What are the dimensions of chatbot service encounters?*”. The answer to this question has demonstrated in the section of 3.5 and the section of 5.3. The section of 3.5 summarized the sources of chatbot service encounters from the previous studies into three dimensions base on their properties (Table 2). These three dimensions are properness of reply, intelligence, and the properness. Then, in the section of 5.3, the sources of different chatbot service encounters from the findings of this study are concluded into the previous three dimensions summarized from the earlier studies (Table 8). Thus, the answer to this question is properness of reply, intelligence, and the properness. However, these three dimensions require more tests and provide a potential research direction for future studies.

## **6.2 Theoretical contributions**

This study contributes to scientific research by making a new opening in the area of technology-mediated service encounters or virtual service encounters, as well as



developing a novel analytic framework to classify the chatbot service incidents. This section describes the main theoretical contributions by connecting the existing literature and the findings of this study.

First of all, this study created a framework (Figure 3) about the chatbot service encounters in OCSEs as the conceptual contribution. This framework covered chatbots, chatbot service encounters, customer satisfaction, the sources of different chatbot service encounters, and consumer behavior, which provides a holistic understanding of the customer experiences literature in the context of chatbot service. Elements in this framework and the relationship between them are supported by the existing studies (Oliver, 1981; Carbonne & Haeckel, 1994; Rose, Hair & Clark, 2011; Klaus, 2013; Michaud, Trevinal & Stenger, 2014; Chan, Barnes & Fukukawa, 2016; Larivière et al., 2017). The model of virtual service quality and customer satisfaction from Seck and Philippe (2013) were partly recognized and supported in this study. Specifically, their findings indicated that the virtual services, ease of use, information quality, and site design influence the virtual service quality and customer satisfaction, which were recognized by this current study. However, this study did not find the security element has an influence on consumers' attitudes toward chatbot service encounters. As well as their model only stated the relationship between these elements with customer satisfaction but did not mention how these factors impact customer satisfaction and the chatbot service encounters. Different from that study, this current study gives a deeper understanding of consumer experiences with chatbot service encounters (how various factors impact consumers' attitude) through analyzing consumers' satisfaction and dissatisfaction incidents with chatbots.

The incident sorting process is another contribution to the methodological value as a novel analytic framework. This incident sorting model with a structure of "decision tree" and consists of two questions, which can divide the chatbot service incidents into three main groups: nature of after-sales services, nature of needs, and nature of unprompted chatbot actions. It was used in this study to sort the critical incidents about chatbot services. This model was inspired by the logic of the incident sorting process from Bitner, Booms and Tetreault's (1990) study about service encounters (offline services). Based on their logic, this study changed the classification process from human-human communication situation to HMC situation, which is helpful for

researchers and managers to classify the chatbot service incidents into different groups. Thus, the incident sorting process can assist the future studies in the chatbot service area and help managers to understand their chatbot services better. Nevertheless, this incident sorting model might be only suitable for the utilitarian text-based chatbots, due to the incidents in this study do not cover the voice-based chatbots, such as Siri from Apple, Google Home from Google, etc. Additionally, another point associated with the methodological value is that the CIT is also a useful and proper method to comprehend chatbot service encounters and chatbot service experiences because it helped the author collected more descriptive and comprehensive data about the chatbot service encounters. This idea supports the Bitner, Booms and Tetreault's (1990) statement that CIT is one of the most appropriate ways to understand customer perceptions.

In accordance with this incident sorting process, this study made another conceptual contribution to the *online customer experience* literature by the findings of the sources of chatbot service encounters. Three main groups of chatbot service encounters with eight sub-group and 16 sources of different types of chatbot service encounters were found. The literature of the customer experience is developing together with technology, from offline- to online. The service encounter was always discussing together with customer satisfaction, which is the result of the customer experience (Oliver, 1981; Rose, Hair & Clark, 2011). The 16 sources found by this study support and complement the ideas from the previous research about chatbot service encounters which has presented in the section of 5.4 (e.g., Mimoun, Poncin & Garnier, 2012; Verhagen et al., 2014, etc.), due to it covered more aspects and situations.

Besides, some substantive contributions which are apart from the research questions but surround the *chatbot services* are identified based on this study. Firstly, this study shows that different goals orient communication in the chatbot situation compared with human-human communication. Communication is goal-driven, and it is mainly about tasks instead of relationships. However, in human-human communication, the goal-driven consists of task-, communication-, and relationship goals (Hobbs & Evans, 1980). Moreover, chatbots become more popular among service areas in these few years. This study found that most consumer needs toward chatbot services are at the basic needs level, i.e., answering their questions instead of performing like a human or

other advanced feature. This situation indicates that chatbot services are not well developed at this moment because consumers do not have high expectations of it. Maslow (1987) stated that people are progressing on higher-level needs after the lower-level needs are satisfied. Thus, the author argues the needs for chatbot services are in the lower-level now, and it points out a direction for future studies.

### **6.3 Managerial implications**

Companies are seeking ways to optimize their chatbot services, enhance their consumer satisfaction, or improve the understanding of different types of chatbot service encounters. This study provides some management ideas regarding what companies could do.

First of all, managers should understand different sources for different chatbot service encounters before developing their chatbot services. Figure 3 provides a straightforward understanding of it. By reading this figure, managers can understand the relationships between chatbots, chatbot service encounters, the sources for different chatbot service encounters, customer satisfaction, and consumer behaviors. Companies can also understand the importance of chatbot service encounters and prevent themselves from the dissatisfactory chatbot service encounters.

Second, managers can develop chatbot service monitor programs by using the incident sorting process established by this study (Figure 2), which could be a useful tool to understand different types of chatbot service encounters. Managers can collect critical incidents from their consumers and put them to the data sorting process, which is able to help them classify the incidents and have a deeper understanding of their consumers' experiences with chatbots. Then, managers are able to get some ideas about how to optimize their chatbot algorithms.

Third, it is crucial to understand consumers' expectations toward chatbot services. This study points out that consumer has different expectations for different types of chatbot services. For chatbots 1) responsible for the basic questions (low-knowledge base), companies should pay attention to train their abilities to identify keywords and prepare concise answers; 2) in charge of the professional questions (high-knowledge base), it

is better to prepare comprehensive solutions which include figure, text, reference time, etc., 3) liable for customized problems, such as based on consumers' data to recommend them clothing style. Chatbots should train (ML) by a variety of permutations of consumers' preferences, and consumers could be co-trainers. For example, a clothing company can organize an event and ask consumers to design the outfit they like to collect data for the machine learning process. Besides, companies should pay more attention to the problem-solving function instead of imitating human behaviors. If chatbots are able to solve consumers' problems, the humanlike behaviors might be the icing on the cake. Otherwise, humanlike behaviors might be useless.

Fourth, combining chatbot with human service is a crucial issue to be considered, especially for the after-sales services (whether consumers or companies cause the problems). It means using machines to augment humans instead of replacing humans, i.e., technologies can strengthen services that support Jarrahi's (2018) and Tripathy's (2018) studies. When chatbots are not able to handle problems or identify keywords, the chatbot could provide a few ideas about what consumers could have a try or give consumers an option to transfer chatbot services to human services. The collaboration between machines and humans for one task can result in better results than singular human services or singular machine services (Collins, 2018).

Fifth, companies ought to understand the purpose of their chatbots' unprompted behaviors. Chatbots pop up with guidance purposes that can help companies to save human labor and help consumers to save time. However, it is better not to use chatbots as a promotion tool because it might change consumer satisfaction to dissatisfaction. For example, a consumer is satisfied with the online purchasing process at the beginning. A chatbot pops up and asks this consumer, "do you want to add something to your card?", the consumer might feel annoyed with this kind of behavior.

Sixth, managers from big companies should pay more attention to build their chatbot services because consumers have higher expectations for chatbots from big companies. Consumers are taking for granted that big companies should have good chatbot services, and they have more dissatisfaction feelings for the poor chatbot services from big companies than SMEs. The poor chatbot services might change consumers'

attitudes toward a brand. However, more research should be done for this point, because in this study, only a part of the participants has this kind of opinion.

#### **6.4 Evaluations of this study: validity and reliability**

The purpose of this section is to interpret the validity and reliability of this study. It briefly introduces the meaning of validity and reliability first, then combines these two points with the content of this study.

On the one hand, validity in qualitative research means “*the precision in which the findings accurately reflect the data*” (Noble & Smith, 2015, p. 34). Loosely speaking, validity is the correctness of the answer (Kirk & Miller, 1986), or to what extent the study measures the original ideas about what the study wants to measure. To improve the validity of the study, using an objective, systematic, and quantified data analysis method is one way. Taking the results back to the participants/interviewees is another way (Brink, 1987).

This study tried to improve that validity from data collection and data analysis perspective. In order to make sure the author can collect the validated data. The author discussed the data collection methods with two professors before applying them, such as the idea of pre-questionnaires, the ways to organize focus group discussions, the suitable sample size, etc. The pre-questionnaires improved the descriptive validity (accuracy of the data) of the data. This was due to its aim to ensure the critical incidents collected by the author were able to meet the requirements and allow participants to understand the research questions. Also, the author asked participants what chatbots did cause their satisfaction and dissatisfaction again at the end of the focus group discussion on purpose to improve the validity of the data. Then, in order to enhance the validity of outcomes, the whole coding process was repeated a few times and was assisted by NVivo. Thus, this objective, systematic, and quantified process helped the author improved the validity.

On the other hand, reliability means the consistency of the analytical procedures (Noble & Smith, 2015, p. 34). Improving reliability is to improve the trustworthiness of one’s study. Loosely speaking, it is about the extent to which the same answer can

be produced in the measurement process (Kirk & Miller, 1986). Using two or more people as raters of the same data under the same decision rules, until an agreement will be created among the raters is a way to improve the reliability of one study (Brink, 1987).

However, one limitation of this study is that the data were analyzed by one person (the author). In order to offset this limitation, the author applied some other ways. The pre-questionnaire plays a supporting role in enhancing the reliability of this study because it provides the author with a pre-understanding of the incidents. Moreover, the data analysis process was taken for a few rounds until there was no more adjustment (sorting, coding, re-sorting, and re-coding). Furthermore, the categories of the data from this study are partly supporting the previous studies, such as most of the ideas from Mimoun, Poncin and Garnier (2012) about the sources of dissatisfactory chatbot service encounters. Also, the dimensions of the source of different chatbot service encounters are initially from the conclusions from existing studies (Table 2). Last but not least, this study was read by two professors and modified by their suggestions, such as providing empirical examples for the results, the name of the framework, etc. Therefore, the validity of this research is guaranteed.

## **6.5 Limitations of the study**

First of all, the limitations regard to the theoretical framework. The previous chapter mentioned that there is no existing definition of chatbot service encounters. Due to chatbot service is a part of the online services, the author modified the concept of OCSE from Klaus (2013) to the chatbot service encounters, it means "consumers' perception of interactions with a chatbot, which causes consumer satisfaction or dissatisfaction". However, this definition requires more studies to verify it. Also, there are different types of chatbots, and this study only focused on the utilitarian text-based chatbots. Thus, it is challenging to define the scope of the literature review.

Moreover, the limitations regard to the data collection and analysis process. On the one hand, for the data collection, this study does not use a proper sample size. Hennink, Kaiser and Weber's (2019) study indicated that the majority points come from the first few groups and the data saturation point in the 6th group. However, due to the time

and cost limitation, this study only organized three focus group discussions. The majority of data was able to be collected, but it was still difficult to ensure data saturation. On the other hand, because of the acknowledged limitation of the data analysis process, the data analysis process was assisted by NVivo, and the author did data sorting and re-sorting a few times. However, the whole process was done by one person, and different coding rules may raise the problems of reliability and validity. Thus, the bias might exist for the data sorting process.

Besides, as this study only focuses on the utilitarian text-based chatbot, the data sorting system developed by this study might not work for other types of chatbots. Different kinds of chatbots might have different functions, such as voice-based chatbots that may have entertainment functions (like a speaker). Thus, the other types of chatbot services might have different data sorting process. Notwithstanding the above, these limitations also provide directions and suggestions for future studies.

## **6.6 Suggestions for future research**

This study provides a deeper understanding of chatbot service encounters and online customer service experience through learning the data collected by the method of CIT and focus group discussions. Based on the data analysis and the limitations listed in the last part, this study provides some directions for future studies.

First and foremost, future research could test the results of this study in a quantitative way, which is helpful in improving the results' credibility of this study. For example, researchers can design a questionnaire by Likert scale to test if consumers agree with the results of this study. The questions could be "I am angry with chatbots to pop up to promote me their products", "I feel satisfied if chatbots provide me a concise answer for a simple question", etc. with the scale from 1-7 (disagree to agree) to test consumers' attitudes.

Moreover, future studies can narrow down the industry range of samples, i.e., focus on one industry to have a deeper understanding. The data collected by this study across all industries make it hard to acquire professional knowledge of a single industry. For example, future studies can replicate the methodology used in this study to the health

care industry and collect the incidents about their chatbot services. Then, the incident sorting process developed by this study can be used to classify the chatbot service incidents and conduct a more focused analysis of the health care industry.

Another derived direction for future studies could be combining consumers' expectations towards chatbot services with Maslow's Hierarchy of Needs theory which suggests how individuals' order of needs are from basic to advanced (physiological, safety, love/belonging, esteem, and self-actualization) (Maslow, 1987) and Kotler's Five Product Level model which presents consumers have different levels of need for products (core-, generic-, expected-, augmented-, and potential product) (Kotler, 2000). The researchers can test consumers' needs for the chatbot services right now, and then comparing it with Maslow's Hierarchy of Needs theory, for example, consumers only need chatbots to answer their questions instead of performing like humans. Consumers' attitudes toward chatbots can be analyzed through using Kotler's Five Product Level model as it can be assumed that consumers only treat chatbot services as an augmented product for human services. This kind of study can complement the existing literature about marketing and assist companies in setting directions for their chatbot services.



## REFERENCES

- Alshare, F. (2020). Jordanian consumers satisfaction with electrical appliances after-sale services. *Management Science Letters*, 10(9), doi:10.5267/j.msl.2020.2.020
- Anonymous (2001). Know the difference between customer and consumer. *Dairy Herd Management*, 38(3), p. 101.
- Arcand, K. (2017). Can chatbots fully replace humans? Not yet: But great customer experiences will rely on leveraging chatbots for their strengths. (CUSTOMER EXPERIENCE). *CRM Magazine*, 21(6), p. 4.
- Bitner, M., Brown, S. & Meuter, M. (2000). Technology infusion in service encounters. *Journal of the Academy of Marketing Science*, 28(1), pp. 138-149. doi:10.1177/0092070300281013
- Bitner, M., Booms, B. & Tetreault, M. (1990). The Service Encounter: Diagnosing Favorable And Unfavorable. *Journal of Marketing*, 54(1), p. 71. doi:10.2307/1252174
- Bowen, D. E. (1986). Managing customers as human resources in service organizations. *Human Resource Management*, 25(3), pp. 371-383. doi:10.1002/hrm.3930250304
- Bowen, D. E. (2016). The changing role of employees in service theory and practice: An interdisciplinary view. *Human Resource Management Review*, 26(1), pp. 4-13. doi: 10.1016/j.hrmr.2015.09.002
- Bowen, J. T. & Chen, S. (2001). The relationship between customer loyalty and customer satisfaction. *International Journal of Contemporary Hospitality Management*, 13(5), pp. 213-217. doi:10.1108/09596110110395893
- Brandtzaeg, P. & Følstad, A. (2017). Why people use chatbots. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10673, pp. 377-392. doi:10.1007/978-3-319-70284-1\_30
- Brave, S., Nass, C. (2002). Emotion in human-computer interaction. In: Jacko, J.A., Sears, A. (eds.) *The human-computer interaction handbook*. L. Erlbaum Associates Inc, NJ, USA. pp. 81–96.
- Brink, P. J. (1987). Editorial: On Reliability and Validity in Qualitative Research. *Western Journal of Nursing Research*, 9(2), pp. 157-159. doi:10.1177/019394598700900201
- Cambridge dictionary (2020a). Intelligence. Retrieved from: <https://dictionary.cambridge.org/dictionary/english/intelligence>

- Cambridge dictionary (2020b). Proper & behavior. Retrieved from:  
<https://dictionary.cambridge.org/dictionary/english/proper>  
<https://dictionary.cambridge.org/dictionary/english/behavior>
- Cambridge dictionary (2020c). Proper & reply. Retrieved from:  
<https://dictionary.cambridge.org/dictionary/english/proper>  
<https://dictionary.cambridge.org/dictionary/english/reply>
- Caruana, A. (2002). Service loyalty: The effects of service quality and the mediating role of customer satisfaction. *European Journal of Marketing*, 36(7/8). Pp. 811-828. DOI: 10.1108/03090560210430818
- Chaffey, D. (2016). Digital marketing (6th ed). *Edinburgh Gate*, Harlow: Pearson Education Limited.
- Ciechanowski, L., Przegalinska, A., Magnuski, M. & Gloor, P. (2019). In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems*, 92, pp. 539-548. doi:10.1016/j.future.2018.01.055
- Clarke, V. & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology: Qualitative Positive Psychology*. Edited by Kate Hefferon and Arabella Ashfield, 12(3), pp. 297-298. doi:10.1080/17439760.2016.1262613
- Chan, K. W., Yim, C. K. (. & Lam, S. S. (2010). Is Customer Participation in Value Creation a Double-Edged Sword? Evidence from Professional Financial Services Across Cultures. *Journal of Marketing*, 74(3), pp. 48-64. doi:10.1509/jmkg.74.3.48
- Chan, S. F., Barnes, B. R. & Fukukawa, K. (2016). Consumer control, dependency and satisfaction with online service. *Asia Pacific Journal of Marketing and Logistics*, 28(4), pp. 594-615. doi:10.1108/APJML-09-2015-0134
- Chung, M., Ko, E., Joung, H. & Kim, S. J. (2018). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research*. doi:10.1016/j.jbusres.2018.10.004
- Collins, R. (2018, January 24). The human-machine symbiosis. *Management issues*. Retrieved from: <https://www.management-issues.com/opinion/7298/the-human-mach-in-symbiosis/>
- Craig, R. T. (1999). Communication Theory as a Field. *Communication Theory*, 9(2), pp. 119-161. doi:10.1111/j.1468-2885.1999.tb00355.x
- Cronin, J. J. & Taylor, S. A. (1992). Measuring service quality: a re-examination and extension. *Journal of Marketing*, 56(3), pp.55–68
- Cyr, D. (2008). Modeling Web Site Design Across Cultures: Relationships to Trust, Satisfaction, and E-Loyalty. *Journal of Management Information Systems*, 24(4), pp. 47-72. doi:10.2753/MIS0742-1222240402

- Dal Porto, L. (2017). Chatbot chatter.(use of chatbots to boost customer service)(The Progress Reports). *Quality Progress*, 50(7), p. 6.
- Dale, R. (2016). The Return of the Chatbots. *Natural Language Engineering*, 22(5), 811–817. DOI: 10.1017/S1351324916000243
- Davis, F.D. (1986), “A technology acceptance model for empirically testing new end-user information systems: theory and results”, doctoral dissertation, *MIT Sloan School of Management*, Cambridge, MA.
- Devaney, E. (2018, January 23). The state of Chatbots Report: How Chatbots Are Reshaping Online Experiences. *Drift*. Retrieved from <https://www.drift.com/blog/Chatbots-report/>
- Diekroger, E. A. (2014). The power of qualitative research. *Pediatrics*, 134(4), p. e933. doi:10.1542/peds.2014-0549
- Feine, J., Morana, S. & Gnewuch, U. (2019). Measuring Service Encounter Satisfaction with Customer Service Chatbots using Sentiment Analysis, *In Proceedings of the 14th International Conference on Wirtschaftsinformatik (WI2019)*, Siegen, Germany.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327–358. doi:10.1037/h0061470
- Forsythe, S., Liu, C., Shannon, D. & Gardner, L. C. (2006). Development of a scale to measure the perceived benefits and risks of online shopping. *Journal of Interactive Marketing*, 20(2), pp. 55-75. doi:10.1002/dir.20061
- Frey, C. B. & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting & Social Change*, 114(C), pp. 254-280. doi:10.1016/j.techfore.2016.08.019
- Garber, M. (2014, June 2014). Chili’s Has Installed More Than 45,000 Tablets in Its Restaurants. *The Atlantic*. Retrieved from <http://www.theatlantic.com/technology/archive/2014/06/chilis-is-installing-tablet-ordering-at-all-its-restaurants/372836/>
- Getrich, C. M., Bennett, A. M., Sussman, A. L., Solares, A. & Helitzer, D. L. (2016). Viewing Focus Groups Through a Critical Incident Lens. *Qualitative Health Research*, 26(6), pp. 750-762. doi:10.1177/1049732315579178
- Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K. & Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human - Computer Studies*, 67(10), pp. 842-849. doi:10.1016/j.ijhcs.2009.07.001
- Gupta, S. & Zeithaml, V. (2006). Customer Metrics and Their Impact on Financial Performance. *Marketing Science*, 25(6), pp. 718-739. doi:10.1287/mksc.1060.0221

- Guzman, A. L. & Lewis, S. C. (2020). Artificial intelligence and communication: A Human–Machine Communication research agenda. *New Media & Society*, 22(1), pp. 70-86. doi:10.1177/1461444819858691
- Halstead, D. & Page, T.J. (1992). The effects of satisfaction and complaining behavior on consumer repurchase intentions. *Journal of Consumer Satisfaction*, 5, pp.1-11.
- Hancock, M. E. (2017). The use of WordPress in online focus group studies. (Report). *The Qualitative Report*, 22(8), p. 2273.
- Hennink, M. M., Kaiser, B. N. & Weber, M. B. (2019). What Influences Saturation? Estimating Sample Sizes in Focus Group Research. *Qualitative Health Research*, 29(10), pp. 1483-1496. doi:10.1177/1049732318821692
- Hines, T. (2000). An evaluation of two qualitative methods (focus group interviews and cognitive maps) for conducting research into entrepreneurial decision making. *Qualitative Market Research: An International Journal*, 3(1), pp. 7-16. doi:10.1108/13522750010310406
- Hobbs, J. R. & Evans, D. A. (1980). Conversation as planned behavior. *Cognitive Science*, 4(4), pp. 349-377. doi:10.1207/s15516709cog0404\_2
- Huang, M. & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, 21(2), pp. 155-172. doi:10.1177/1094670517752459
- Hutchinson, A. (2019, May 1). Facebook Messenger by the Numbers 2019. *SocialMediaToday*. Retrieved from: <https://www.socialmediatoday.com/news/facebook-messenger-by-the-numbers-2019-infographic/553809/>
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), pp. 577-586. doi:10.1016/j.bushor.2018.03.007
- Jenkins, M., Churchill, R., Cox, S. & Smith, D. (2007). Analysis of user interaction with service oriented chatbot systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4552(3), pp. 76-83.
- Julia, N. (2018, November 21). E-commerce chatbots: 22 ways to increase sales, conversions & Retention. *Shopifyplus*. Retrieved from <https://www.shopify.com/enterprise/ecommerce-chatbots>
- Kanda, T., Miyashita, T., Osada, T., Haikawa, Y. & Ishiguro, H. (2008). Analysis of Humanoid Appearances in Human-Robot Interaction. *IEEE Transactions on Robotics*, 24(3), pp. 725-735. doi:10.1109/TRO.2008.921566
- Kirk, J., & Miller, M. L. (1986). *Qualitative Research Methods: Reliability and validity in qualitative research*. Newbury Park, CA: SAGE Publications, Inc. doi: 10.4135/9781412985659

- Klaus, P. (2013). The case of Amazon.com: Towards a conceptual framework of online customer service experience (OCSE) using the emerging consensus technique (ECT). *Journal of Services Marketing*, 27(6), pp. 443-457. doi:10.1108/JSM-02-2012-0030
- Klaus, P. (2020). Customer experience, not brands will be on the iron throne. *International Journal of Market Research*, 62(1), pp. 6-8. doi:10.1177/1470785319858570
- Kleinschafer, J., Morrison, M. & Dowell, D. (2018). The importance of the service encounter in influencing identity salience and volunteering behavior in the cultural sector. *International Review on Public and Nonprofit Marketing*, 15(4), pp. 455-474. doi:10.1007/s12208-018-0209-3
- Kotler, P. (2000). *Marketing Management: The Millennium Edition*. Prentice Hall, Inc. Edition: 10th.
- Koufteros, X., Verghese, A., & Lucianetti, L. (2014). The effect of performance measurement systems on firm performance: A cross-sectional and a longitudinal study. *Journal of Operations Management*. 32(6), pp.313-336.
- Labecki, A., Klaus, P., & Zaichkowsky, J. L. (2018). How bots have taken over brand choice decisions. In K. Arai, R. Bhatia & S. Kapoor (Eds.), *Proceedings of the Future Technologies Conference* (pp. 976–989). Cham: Springer.
- Larsson, R., & Bowen, D. E. (1989). Organization and customer: Managing design and coordination of services. *Academy of Management Review*, 14(2), 213–233.
- Larivière, B., Bowen, D., Andreassen, T. & Kunz, W., Sirianni, N., Voss, C., Wunderlich, N. & De Keyser, A. (2017). “Service Encounter 2.0”: An investigation into the roles of technology, employees and customers. *Journal of Business Research*, 79, 238-246. doi: 10.1016/j.jbusres.2017.03.008
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), doi:10.1177/2053951718756684
- Lee, S. M., & Lee, D. (2020). “Untact”: A new customer service strategy in the digital age. *Service Business*, 14(1), pp. 1-22. doi: <http://dx.doi.org/10.1007/s11628-019-00408-2>
- Lee, D., Oh, K. J., & Choi, H. J. (2017, February). *The Chatbot feels you: A counseling service using emotional response generation*. Big data and smart computing (BigComp), *IEEE International Conference*. pp. 437–440.
- Li, Y. (2014). Main factors affecting the online service satisfaction-an empirical study in China. *International Journal of Smart Home*, 8(3), pp. 131-144. doi:10.14257/ijsh.2014.8.3.12

- Maslow, A. H. (1987). *Motivation and personality* (3rd ed.). Delhi, India: Pearson Education.
- Maxham, J. G. & Netemeyer, R. G. (2002). Modeling customer perceptions of complaint handling over time: The effects of perceived justice on satisfaction and intent. *Journal of Retailing*, 78(4), pp. 239-252. doi:10.1016/S0022-4359(02)00100-8
- Mctear, M., Callejas, Z. & Griol, D. (2016). *The Conversational Interface: Talking to Smart Devices*.
- Mead, G. H. (1967), *Mind, Self, & Society: From the Standpoint of a Social Behaviorist*, vol. 1. Chicago, IL: *University of Chicago Press*.
- Mimoun, B. M. S., Poncin, I. & Garnier, M. (2012). Case study—Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer Services*, 19(6), pp. 605-612. doi: 10.1016/j.jretconser.2012.07.006
- Mitchell, T. M. (1997). *Machine learning*. New York (NY): McGraw-Hill.
- Mills, P. K., Chase, R. B. & Margulies, N. (1983). Motivating the client/employee system as a service production strategy. *Academy of management review*. *Academy of Management*, 8(2), p. 301. doi:10.5465/amr.1983.4284740
- Mou, Y. & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior*, 72, pp. 432-440. doi:10.1016/j.chb.2017.02.067
- Noble, H. & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence Based Nursing*, 18(2), p. 34. doi:10.1136/eb-2015-102054
- Noone, B. M., Kimes, S. E., Mattila, A. S., & Wirtz, J. (2009). Perceived service encounter pace and customer satisfaction: An empirical study of restaurant experiences. *Journal of Service Management*, 20(4), 380-403. doi: <http://dx.doi.org/10.1108/09564230910978494>
- Oliver, R. L. (1981). Measurement and Evaluation of Satisfaction Processes in Retail Settings. *Journal of Retailing*, 57(3), p. 25.
- Oliver, R. L. (1997). *Satisfaction: A behavioral perspective on the consumer*. McGraw Hill, New York
- Ong, B. K. (2012). Grounded Theory Method (GTM) and the Abductive Research Strategy (ARS): A critical analysis of their differences. *International Journal of Social Research Methodology*, 15(5), pp. 417-432. doi:10.1080/13645579.2011.607003
- Patrício, L., Fisk, R. P., Falcão E Cunha, J. & Constantine, L. (2011). Multilevel Service Design: From Customer Value Constellation to Service Experience

- Blueprinting. *Journal of Service Research*, 14(2), pp. 180-200. doi:10.1177/1094670511401901
- Polit, D. F. & Beck, C. T. (2010). Generalization in quantitative and qualitative research: Myths and strategies. *International Journal of Nursing Studies*, 47(11), pp. 1451-1458. doi:10.1016/j.ijnurstu.2010.06.004
- Prahalad, C. & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing*, 18(3), pp. 5-14. doi:10.1002/dir.20015
- Ranjan, A. & Mulakaluri, N. (2018). Leveraging Chatbots for 'the Edge' in Business. *Gavesana Journal of Management*, 10(1), pp. 1-5.
- Rainie, L. & Anderson, J. (2017, June 6). *Internet & Technology*: The internet of things connectivity binge: what are the implications? Retrieved from: <https://www.pewresearch.org/internet/2017/06/06/the-internet-of-things-connectivity-binge-what-are-the-implications/>
- Resnick, S., Foster, C. & Woodall, T. (2014). Exploring the UK high street retail experience: Is the service encounter still valued? *International Journal of Retail & Distribution Management*, 42(9), pp. 839-859. doi:10.1108/IJRDM-05-2013-0090
- Reynolds, K. E. & Beatty, S. E. (1999). A relationship customer typology. *Journal of Retailing*, 75(4), pp. 509-523. doi:10.1016/S0022-4359(99)00016-0
- Rose, S., Hair, N. & Clark, M. (2011). Online Customer Experience: A Review of the Business-to-Consumer Online Purchase Context. *International Journal of Management Reviews*, 13(1), pp. 24-39. doi:10.1111/j.1468-2370.2010.00280.x
- Sandry, E. (2018). *Encounter, story and dance: Human-machine communication and the design of human-technology interactions*.
- Seck, A. M. & Philippe, J. (2013). Service encounter in multi-channel distribution context: Virtual and face-to-face interactions and consumer satisfaction. *The Service Industries Journal*, 33(6), pp. 565-579. doi:10.1080/02642069.2011.622370
- Shobeiri, S., Mazaheri, E., & Laroche, M. (2018). Creating the right customer experience online: The influence of culture. *Journal of Marketing Communications*, 24(3), 270–290.
- Smith, A. K., Bolton, R. N. & Wagner, J. (1999). A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research*, 36(3), pp. 356-372. doi:10.2307/3152082
- Skowron, M., Rank, S., Theunis, M., Sienkiewicz, J. (2011). The Good, the Bad and the Neutral: Affective Profile in Dialog System-User Communication. In: D’Mello, S., Graesser, A., Schuller, B., Martin, J.-C. (eds.) *Affective Computing and Intelligent Interaction*, pp. 337–346. Springer, Berlin, Heidelberg.

- Stewart, D. W. & Shamdasani, P. (2017). Online Focus Groups. *Journal of Advertising: Themed Issue - Methodology in Advertising Research*, 46(1), pp. 48-60. doi:10.1080/00913367.2016.1252288
- Surprenant, C. & Solomon, M. (1987). Predictability and Personalization in the Service Encounter. *Journal of Marketing*, 51(2), p. 86. doi:10.2307/1251131
- Söderlund, M. (2017). Employee display of burnout in the service encounter and its impact on customer satisfaction. *Journal of Retailing and Consumer Services*, 37(C), pp. 168-176. doi:10.1016/j.jretconser.2016.10.014
- Tax, S. S., Mccutcheon, D. & Wilkinson, I. F. (2013). The Service Delivery Network (SDN): A Customer-Centric Perspective of the Customer Journey. *Journal of Service Research*, 16(4), pp. 454-470. doi:10.1177/1094670513481108
- Taylor, S. (1994). Waiting for service: The relationship between delays and ev. *Journal of Marketing*, 58(2), pp. 56-69. doi:10.2307/1252269
- Tenney, M. (2016). Tips for Holding an Effective Online Focus Group. *Nonprofit Communications Report*, 14(4), p. 4. doi:10.1002/npcr.30386
- Trevinal, M. A. & Stenger, T. (2014). Toward a conceptualization of the online shopping experience. *Journal of Retailing and Consumer Services*, 21(3), pp. 314-326. doi: 10.1016/j.jretconser.2014.02.009
- Tripathy, A. (2018). Why Humans Need To Find Symbiosis With AI. *Forbes*. Retrieved from: <https://www.forbes.com/sites/forbestechcouncil/2018/01/29/why-humans-need-to-find-symbiosis-with-ai/#6c407d35407a>
- Trotter, C. (2017, May 1). Retail Trends Factfile 2017: Physical Retail. *Insider Trends*. Retrieved from <http://www.insider-trends.com/retail-trends-factfile-2017-physical-retail>
- Tse, D. K., Wilton P. C. (1988) Models of consumer satisfaction formation: an extension. *Journal of Marketing Research*. 25(2), pp. 204–212
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), pp. 433-460.
- Van Doorn, J. & Verhoef, P. (2008). Critical incidents and the impact of satisfaction on customer share. *Journal of Marketing*, 72(4), pp. 123-142. doi:10.1509/jmkg.72.4.123
- Verhagen, T., Van Nes, J., Feldberg, F. & Van Dolen, W. (2014). Virtual Customer Service Agents: Using Social Presence and Personalization to Shape Online Service Encounters. *Journal of Computer - Mediated Communication*, 19(3), pp. 529-545. doi:10.1111/jcc4.12066



- Viergever, R. F. (2019). The Critical Incident Technique: Method or Methodology? *Qualitative Health Research*, 29(7), pp. 1065-1079. doi:10.1177/1049732318813112
- Wakefield, J. (2016, March 24). Microsoft chatbot is taught to swear on Twitter. *BBC News*. Retrieved from: <https://www.bbc.com/news/technology-35890188>
- Wallin Andreassen, T. (2000). Antecedents to satisfaction with service recovery. *European Journal of Marketing*, 34(1/2), pp. 156-175. doi:10.1108/03090560010306269
- Wan, L. C., & Chan, E. (2019). Failure is Not Fatal: Actionable Insights on Service Failure and Recovery for the Hospitality Industry. *Boston Hospitality Review*. Retrieved from: <https://www.bu.edu/bhr/files/2019/02/Failure-is-Not-Fatal-Actionable-Insights-on-Service-Failure-and-Recovery-for-the-Hospitality-Industry-Lisa-Wan-and-Elisa-Chan.pdf>
- Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S. & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, 29(5), pp. 907-931. doi:10.1108/JOSM-04-2018-0119
- Wolfinbarger, M. F. & Gilly, M. (2003), “eTailQ: dimensionalizing, measuring, and predicting e-Tail quality”, *Journal of Retailing*, 79(3), pp. 183-198.
- Wu, C., Huang, K., Fiegantara, S. & Wu, P. (2012). The impact of online customer satisfaction on the yahoo auction in Taiwan. *Service Business*, 6(4), pp. 473-487. doi:10.1007/s11628-012-0155-9
- Wunderlich, N. & Paluch, S. (2018). *A Nice and Friendly Chat with a Bot: User Perceptions of AI-Based Service Agents*.
- Yi, Y. & Gong T. (2008). The effects of customer justice perception and affect on customer citizenship behavior and customer dysfunctional behavior. *Industrial Marketing Management*. 37(7), pp. 763-783.
- Yin, R.K. (1994), *Case Study Research: Design and Methods*, 2nd ed., Sage Publications, London.
- Zhang, W., Liu, T., Wang, Y., & Zhu, Q. (2017). Neural personalized response generation as domain adaptation. (arXiv preprint arXiv:1701.02073).
- Zhu, P., Zhang, Z., Li, J., Huang, Y., & Hai, Z. (2018). Lingke: a Fine-grained Multi-turn Chatbot for Customer Service. *ArXiv*, abs/1808.03430.

## APPENDICES

### Appendix 1

#### PRE-QUESTIONNAIRE

Hi, thank you so much for joining my study!

I am Xinyi, a master's student in Marketing at the University of Oulu. I'm doing my master thesis about chatbots and consumer experiences. Nowadays, many companies have started to use chatbots to replace human services. Thus, it is important to have you join my study and help me to understand the existing chatbot services and optimize them. I will provide chocolate to express my gratitude and would like to be a helping hand for you when you collect data for your thesis later.

Please think of two specific experiences when you had memorable experiences with a text-based chatbot who was representing a shop/firm. One is the most favorable and one is the most unfavorable experience. Please write down the whole interactive process, including 1) For what types of industry/product. 2) The reason why you chat with chatbots. 3) What are the chatbots' characteristic (Like a human? Has a name? Has a gender? etc.). 4) How's the interaction went. 5) What the chatbot did made you feel good/bad? 6) Could you describe your feelings and emotions? 7) What you did after the interaction (Sharing your experiences with your friends? Has your attitude towards this brand changed? etc.). Please write down the favorable and unfavorable experiences separately.

Notice: The interaction should be with a machine instead of a real human, it should be the whole experience, and it should be from your perspective (consumer).

1. Please describe the most favorable experience with a chatbot, I would like you to describe the experience with at least 120 words.
2. Please describe the most unfavorable experience with a chatbot, I would like you to describe the experience with at least 120 words.

**Appendix 2****SEMI-STRUCTURED FOCUS GROUP DISCUSSION**

Tips: if the answers mentioned in the individual sharing section already, then the comparable question be skipped.

- Does this chatbot have specific characteristics? Female/Male/No particular? Have a name? Have a profile photo? In which language? Friendly or not?
- How did the chatbot reply to you? Like a machine or a friend?
- What is your overall evaluation of this experience?
- In your opinion, what is the fuse made you satisfy/dissatisfy for this chatbot service?
- What did you do after the favorable and unfavorable situations happened? Such as give them bad comments online, sharing with your friends, etc.
- To what extend did the chatbot meet your expectations? Is there a gap? Which kind of gap?
- Have you changed your opinions about this brand because of the chatbot?
- Do you mind if chatbots use your data to adjust its characteristics? (customize the services)
- Do you want to keep friendship with chatbots? Let them keep your chat history and able to continue the topic from last time?
- Comparing with human services, do you think chatbot is better? Why?
- Will you trust this chatbot later after the favorable/unfavorable encounter happen?