

## When and How to Implement Choices on Customer Service Chatbots

Elizabeth Han  
 McGill University  
 elizabeth.han@mcgill.ca

Dezhi Yin  
 University of South Florida  
 dezhiyin@usf.edu

Han Zhang  
 Georgia Institute of Technology  
 han.zhang@scheller.gatech.edu

### Abstract

*Many service chatbots are equipped to provide choices when interacting with customers to streamline the service delivery process. This research investigates when and why the implementation of choices enhances or impairs customers' service experience. Based on the concept of fluency, we posit that the choice implementation is beneficial only after a conversational breakdown due to a chatbot failure; otherwise, the value of choice provision for facilitating fluency may not be salient enough. We further propose that choice provision is counterproductive when the choice set is incomprehensive, reducing (rather than enhancing) the fluency in the use of provided choices for a subsequent decision. We conducted several experiments to test these hypotheses. By illuminating when and why choice implementation may help or harm customers during a chatbot-initiated service interaction, we augment the current understanding of a chatbot's role in customers' service experience and provide insights for the deployment of choice-equipped service chatbots.*

**Keywords:** service chatbot, choice, conversational breakdown, choice comprehensiveness, fluency.

### 1. Introduction

Firms are increasingly adopting artificial intelligence (AI) powered applications to streamline various business processes. Especially, a chatbot—a text-based conversational agent—is projected to reach a global market size of over \$100 billion by 2026 (Mordor Intelligence 2022). Chatbots have been widely deployed in customer service to provide a wide range of service tasks, from answering simple questions to giving recommendations and advice (Markets and Markets 2019). The implementation of service chatbots increases the efficiency of service delivery processes and reduces the costs of both physical and emotional labor for frontline employees. Thus, the role of chatbots in customer service, especially the way they communicate with customers,

has been vital for both researchers and practitioners (Crolic et al. 2022; Fotheringham and Wiles 2022; Huang and Rust 2021; Luo et al. 2019).

Although chatbots are developed based on natural language processing (NLP) technologies and are capable of understanding and speaking human languages to a certain extent, such technologies are far from perfect (Ashktorab et al. 2019; Benner et al. 2021). When customers interact with a chatbot, they often expect the technology-induced service to provide a smooth and seamless experience (Ostrom et al. 2021). To prevent or reduce the likelihood of breakdowns that usually happen to chatbots, firms often implement choices—structured message templates in the form of a guided conversation, during which customers can select one of the provided options as their input message (Klopfenstein et al. 2017). Equipping chatbots with choices is deemed especially suitable for routine and standardized service tasks (Huang and Rust 2021).

Although the implementation of choices in chatbots is prevalent in the industry, there is not much empirical evidence for its impact on service outcomes. Human-based service interactions tend to involve natural conversations, so providing choices in this process may be unnecessary and break the natural flow of conversations. Therefore, choice provision is a unique characteristic of chatbot-based service interactions. While the popularity of choice-equipped chatbots indicates a generally favorable view of this feature, choice implementation may also have unintended consequences under certain situations.

To shed light on the impact of this crucial practice, we examine when the implementation of chatbot-initiated choices during a service interaction is beneficial or counterproductive. The primary reason for implementing choices in practice is to enhance the fluency of a service process. Fluency refers to the ease of processing ongoing tasks or information (Oppenheimer 2008). Because customers' perception of fluency influences service outcomes (Fernández-Sabiote and López-López 2020), we focus on two variables that are especially relevant to the fluency of chatbot-initiated service interactions: conversational

breakdowns and choice comprehensiveness. Conversational breakdowns refer to breakdowns that occur due to chatbots' imperfect capability to understand users' messages (Ashktorab et al. 2019). Choice comprehensiveness refers to the extent to which a choice set includes all relevant information (Forbes 2007). We tested the role of these boundary conditions on the impact of choice using a series of experimental studies. The studies provided consistent evidence supporting our hypotheses. Theoretical and practical implications are also discussed.

## 2. Theoretical development

### 2.1 Choice-providing chatbots

Many firms adopt chatbots to be at the frontline of interacting with customers (Kulkarni 2023). To facilitate a social and interpersonal environment, recent research has emphasized the need for equipping chatbots with the ability to have natural and human-like conversations (Fotheringham and Wiles 2022; Huang and Rust 2021; Schanke et al. 2021). However, using NLP technologies to emulate human often backfires due to inauthenticity, reduced perception of control over a chatbot, uncanniness, and violation of expectations (Fotheringham and Wiles 2022; Han et al. Forthcoming; Kim et al. 2019; Nguyen et al. 2022). Moreover, during technology-based service encounters, customers tend to seek quick, efficient, and task-oriented interactions (Meuter et al. 2000). For efficiency and speed purposes, practitioners often equip chatbots with choices from which users can choose to generate a guided conversation (Klopfenstein et al. 2017). This feature is popularly deployed to deal with routine, systematic tasks without incurring much cost of applying a more advanced technology (Klopfenstein et al. 2017).

A chatbot's provision of choices indeed enables users to make a quick response, increases efficiency, and minimizes the risk of errors. Moreover, the provision of choices reduces users' cognitive load and increases the perception of autonomy by allowing them to select their own action, as often achieved through menu-based interfaces of a traditional website and self-service technologies (Meuter et al. 2000; Nguyen et al. 2022). While these advantages are recognized by practitioners, there are not much empirical evidence on the impact of a choice-providing chatbot. Its advantages are undoubted, but it is possible that such advantages are only realized in certain situations. Furthermore, providing choices may be counterproductive depending on how it is presented. Thus, we aim to illuminate the potential boundary conditions for such an impact.

### 2.2 The role of fluency in service interactions

A primary reason for implementing choices in a service chatbot is to achieve frictionless interaction with customers. Furthermore, users typically expect technology-induced service encounters to provide seamless interactions (Bitner et al. 2000). Processing fluency, which refers to a subjective experience of how information or a task is easily processed, is known to significantly impact people's judgments and decision-making (Alter and Oppenheimer 2009). Frequently, an individual's judgment based on certain stimuli is determined by how quickly and easily the stimuli are processed rather than the stimuli themselves (Schwarz et al. 1991).

Prior literature on customer service has also acknowledged the importance of fluency for crucial business outcomes, such as brand evaluation and service evaluation (Fernández-Sabiote and López-López 2020; Shen et al. 2018). While fluency examined in the customer service and general fluency literatures takes various forms (e.g., visual, linguistic, semantic), the underlying notion is that the perception of fluency is driven by the ease and the speed of processing stimuli (Alter and Oppenheimer 2009). Along the similar lines, a chatbot's provision of choices can facilitate fluency as it allows customers to respond quickly and enhance the ease of proceeding with the ongoing interaction and the service task. This increased fluency is known to generate positive affective responses and cognitive judgments (Winkielman et al. 2003).

However, the supposed positive effect may not always materialize because the perception of fluency may vary based on situational factors, such as when and how choices are presented. Therefore, we focus on two relevant boundary conditions for the impact of choices: conversational breakdowns and choice comprehensiveness. First, during a service interaction that is already flowing well, the effect of choices on the perception of fluency may not be salient enough. Instead, the value of choices is more likely to be recognized during an interaction experiencing disruptions. A common disruption in our context is a conversational breakdown due to chatbot failures (Ashktorab et al. 2019). For instance, when a conversational breakdown occurs, providing choices may restore the lost perception of fluency. Second, the structure of a choice set can also matter, similar to the role of choice architecture in rational decision-making (Thaler and Sunstein 2008). For instance, if the choice set does not include everything related to a customer's needs or requests (e.g., when a customer makes a complex or unique request), providing choices may

undermine fluency. Next, we explain each of these boundary conditions and present our hypotheses.

### 2.3 Choice implementation after conversational breakdowns

Due to the imperfect NLP technology and increasing sophistication in users' requests, high failure rates are often observed for chatbots (Ashktorab et al. 2019; Simonite 2017). Although the recent large language models (LLM) may alter the maturity of chatbots, still many firms are yet to actively deploy LLM-based service chatbots in practice. Acknowledging the imperfection of AI technologies, several studies have investigated the impact of AI failures, but they provided mixed evidence (Sheehan et al. 2020; Choi et al. 2020).

One of the most prevalent chatbot failure is its inability to understand a user's message (Ashktorab et al. 2019). Such failures can be technical errors due to an inherent problem with the chatbot's system or interaction failures due to "incomprehensible" user messages beyond the chatbot's capability. As malfunctioning of an IT system may disrupt an individual's flow of work (Addas and Pinsonneault 2015), service chatbot failures will incur a conversational breakdown, disrupting the flow of the service delivery process. During a traditional human-delivered service, errors made in the process create interruptions and inhibit the fluent service delivery process (Froehle and White 2014; Seshadri and Shapira 2001; Stewart and Chase 1999). For a chatbot, a prominent source of disfluency comes from its inability to understand a message from humans. Such a breakdown can trigger a customer to anticipate a potential service failure, consequently impairing the perception of fluent service delivery.

We argue that the choice provision is more likely to benefit service outcomes when conversational breakdowns occur. Because providing choices can guide customers back to a structured conversation, it can repair the disruption from the breakdowns. Specifically, when conversational breakdowns cause a salient disruption to a service process, the provision of choices can prevent future disruptions and improve the perception of fluency. The perception of fluency can be also linked to the perception of the chatbot's competence, which is associated with traits that portray one's ability (e.g., capability, efficacy) (Fiske et al. 2007), as fluent service interactions are generally deemed as a primary capability of technology-mediated service interaction (Meuter et al. 2000).

The increased perception of the chatbot's competence due to its act of restoring fluency (i.e., providing choices) can ultimately enhance customers'

service evaluations. Indeed, prior customer service literature has provided extensive evidence on the impact of the perception of service employees' competence on service evaluations (Parasuraman et al. 1985). In this research, we focus on customers' perceptions of a service encounter because of their implications on various business outcomes (Heskett et al. 1994). In sum, we propose the following:

**Hypothesis 1:** The choices provided by a chatbot after conversational breakdowns enhance service evaluations. In contrast, such an effect disappears when there are no conversational breakdowns.

**Hypothesis 2:** The positive effect of chatbot-provided choices after conversational breakdowns is driven by a customer's increased perception of fluency, which enhances the perception of the chatbot's competence.

### 2.4 The comprehensiveness of a choice set

While a conversational breakdown serves as a contextual factor that makes the perception of fluency salient, thus enabling the effect of providing choices to materialize, the characteristics of a choice set may also drive whether the provided choice set enhances fluency or not. Indeed, during rational decision-making, the structure of a choice set plays a significant role. The design of a choice set is also known as choice architecture, which is a vital driving force that can nudge a decision maker's behavior (Thaler and Sunstein 2008). Several elements compose choice architecture, such as the number, descriptions, and presentation format of choices (Johnson et al. 2012). Among various elements, one element that is relevant to chatbot-provided choices and commonly observed in practice is choice comprehensiveness.

Comprehensiveness is a construct commonly used in strategic decision-making literature, which refers to how extensively one gathers information when making strategic decisions (Fredrickson 1984). Comprehensiveness is especially important in unstable and uncertain environments which require the consideration of all relevant information, including determinate ones, for enhancing decision quality (Forbes 2007). We extend this construct primarily studied at an organizational level to the context of individual decision-making. Similar to the definition of comprehensiveness from the prior literature, we define choice comprehensiveness as the extent to which a choice set provides an exhaustive list of information, including determinate ones.

In our context, when a chatbot provides choices, ensuring comprehensiveness of the choice set is critical to allow individuals to make an appropriate decision. Especially after a conversational breakdown, which may increase customers' perception of

environmental instability and uncertainty, the comprehensiveness of a choice set will matter even more. However, achieving choice comprehensiveness is often less feasible in practice. Choices are generally predefined in a system before an interaction begins and are not flexible enough to be changed during the interaction. Thus, from a practitioner's standpoint, it is essential to design choices that can satisfy every customer. However, it is unlikely that a predefined choice set with a limited number of options can encompass every potential request. For instance, a service task with high complexity often requires service technologies to personalize and be aware of specific demands (Xu et al. 2014), but inflexible, predefined choices will not be able to accommodate every personal need. Also, a customer may make ambiguous and uncertain demands, where there is a high chance that a customer does not find a satisfying option from the predefined choice set. If a choice set is not comprehensive, especially after a conversational breakdown, it will further undermine the perception of fluency by causing difficulty in decision-making (Alter and Oppenheimer 2009).

Moreover, how fluently an individual decides his or her subsequent action depends on the alignment of that action and any stimuli provided right before deciding to commit to that action (Chambon and Haggard 2012). Similarly, the extent to which a customer's anticipated action aligns with the provided choices may drive the fluency of the choice process. An incomprehensive choice set will deter fluency because none of the choices align with a customer's expected action. Altogether, the decreased perception of fluency will undermine the perception of the chatbot's competence as the disrupted fluency from an incomprehensive choice set can be attributed to the chatbot's capability. Such decrease in perceived competence will drive a negative perception of an entire service experience. Thus, we propose:

**Hypothesis 3:** The positive effect of providing choice after conversational breakdowns backfires when the choice set is incomprehensive, thus hurting service evaluations even more than in the absence of choice.

**Hypothesis 4:** The negative effect of an incomprehensive choice set provided after conversational breakdowns is driven by a customer's decreased perception of fluency, which reduces the perception of the chatbot's competence.

We investigated our hypotheses through a series of experimental studies in which participants interacted with a service chatbot to resolve a service issue based on a hypothetical scenario. We introduce the first study in the next section.

### 3. Study 1

The goal of Study 1 was to investigate whether a conversational breakdown moderates the impact of a chatbot's choice provision on service evaluations as well as the role of customers' perceptions of fluency and the chatbot's competence, as proposed in Hypotheses 1 and 2. To do so, we manipulated the presence of choices and conversational breakdowns in a 2 X 2, between-subjects design.

#### 3.1. Stimulus materials

We used a predesigned script for the chatbot's messages to ensure that every aspect of the interaction remains identical, except for the presence of choice and conversational breakdown. To manipulate the presence of conversational breakdown, we inserted error messages that the chatbot could not understand the participant's message. We used such error messages because not understanding a customer's message is one of the most common pitfalls of a chatbot that disrupts conversational flow (Benner et al. 2021). While those in the breakdown conditions encountered several error messages throughout the chat, those in the no-breakdown conditions did not encounter any error messages. We slightly varied each of the inserted error messages, but overall, these error messages asked participants to rephrase what they had said right before. Then, the chat continued as in the no-breakdown conditions.

We manipulated the presence of choice by varying whether participants freely type in their messages or click and choose one of the options provided by the chatbot as their message. For instance, among the no-breakdown conditions, when the chatbot asked participants to describe a service issue, those in the no-choice condition would type in their response and then see the subsequent message from the chatbot. To those in the choice condition, the chatbot provided three options: 'Missing item,' 'Check order status,' and 'Return/exchange item(s).' The participants could see the subsequent message from the chatbot only after they chose one of the options. Also, the chatbot provided choices every time. Meanwhile, in the breakdown, choice condition, choices were provided only after a conversational breakdown. In the breakdown, no-choice condition, the chatbot did not provide any choices after a breakdown, and participants had to type in their messages.

## 3.2 Procedures and measures

Three hundred and thirty-nine undergraduate students (188 female) from a U.S. university participated in the study in exchange for course credit. Participants were randomly assigned to one of the four conditions: breakdown or no-breakdown and choice or no-choice.

Participants first encountered the cover story that described a recent delivery from an online clothing store in which one of the items was missing. We chose the online retail store as the setting because virtual chats are commonly deployed in this industry to communicate with customers. Furthermore, chatbots are increasingly becoming the first point of contact for service in this industry (Kulkarni 2023). For the service-related issue, we chose a standardized service task for practical and design reasons: first, most chatbots are deployed to handle standardized service tasks in practice, and second, using such a standardized task reduces the risk of a chatbot making inconsistent responses to participants and ensures procedure equivalence across conditions.

During the chat, as each message from the chatbot appeared, participants had to type in their response underneath or choose one of the options provided by the chatbot depending on their experimental condition before seeing the subsequent message. Throughout the chat, participants saw a reminder of the critical facts from the script next to the chat interface so that they would not forget the key details.

After the chat, participants evaluated the service by reporting their perception of service quality and satisfaction with the service. Customers' perception of service quality is critical for service providers because it is an overall evaluation of service outcome, interaction, and environment associated with vital organizational outcomes (Brady and Cronin 2001). Perceived service quality was measured using three items (e.g., "poor / excellent"). Customers' satisfaction with the service is also essential as it is a key predictor of their intention to continue using the service (Oliva et al. 1992). Satisfaction was measured using three questions (e.g., "how satisfied or dissatisfied did your experience with the service agent leave you feeling?"). Both were measured on a seven-point semantic differential scale and adapted from Brady and Cronin (2001). Participants also answered two attention check questions, which were later used to ensure subject quality. In addition, we measured participants' perception of fluency of their service experience using six items (e.g., "flowing very unwell / flowing very well") (Graf et al. 2018) and perception of the chatbot's competence using six items (e.g., "not at all capable / extremely capable") (Fiske et al. 2007)

on a seven-point semantic differential scale. As manipulation checks, participants were asked how often they clicked and chose from options provided during the chat and how often they encountered interruptions on a five-point scale ('never' equals 1; 'always' equals 5) (Speier et al. 1999).

## 3.3 Results

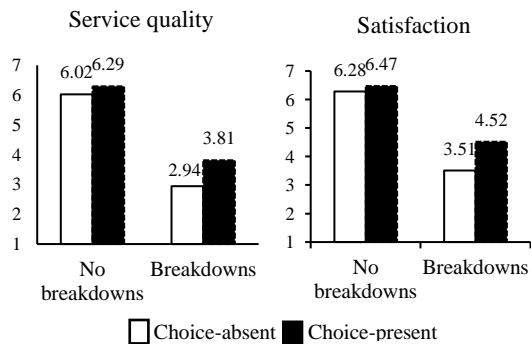
In our analyses, we used the responses from 303 subjects who passed the two attention checks. We confirmed our manipulations by observing that participants in the choice conditions perceived that they had to choose options more often than those in the no-choice conditions ( $M_{choice} = 4.24$  vs.  $M_{no-choice} = 1.56$ ,  $SDs = .94$  and  $.96$ ,  $t(301) = 24.660$ ,  $p < .001$ ), and participants in the breakdown conditions perceived the breakdowns to have occurred more frequently than those in the no-breakdown conditions ( $M_{breakdown} = 2.31$  vs.  $M_{no-breakdown} = 1.22$ ,  $SDs = 1.25$  and  $.69$ ,  $t(301) = 9.433$ ,  $p < .001$ ).

Next, to test the moderating impact of conversational breakdowns, we conducted a two-way ANOVA with the presences of choice and conversational breakdown as two between-subjects factors, and perceived service quality and satisfaction with the service as two outcome variables. We found a significant interaction effect of choice and conversational breakdown on perceived service quality ( $F(1,299) = 3.987$ ,  $p = .047$ ) and on satisfaction ( $F(1,299) = 8.803$ ,  $p = .003$ ). Pairwise comparisons further showed that when there were no conversational breakdowns, providing choices did not have any significant effect on either perceived service quality ( $M_{choice} = 6.29$  vs.  $M_{no-choice} = 6.02$ ,  $F(1,299) = 1.600$ ,  $p = .2$ ) or satisfaction ( $M_{choice} = 6.47$  vs.  $M_{no-choice} = 6.28$ ,  $F(1,299) = .984$ ,  $p = .3$ ). On the other hand, when there were conversational breakdowns, providing choices significantly enhanced the perception of service quality ( $M_{choice} = 3.81$  vs.  $M_{no-choice} = 2.94$ ,  $F(1,299) = 16.495$ ,  $p < .001$ ) and satisfaction ( $M_{choice} = 4.52$  vs.  $M_{no-choice} = 3.51$ ,  $F(1,299) = 26.592$ ,  $p < .001$ ). These findings indicate that a chatbot's provision of choices benefits customers' service experience only after conversational breakdowns, thus confirming Hypothesis 1. Figure 1 depicts the interactions.

To examine the underlying mechanism for the observed interaction, we first tested the moderating role of conversational breakdown for the impact of the choice provision on perceived fluency. We confirmed a significant interaction effect of the presences of conversational breakdown and choice on the perception of fluency ( $F(1,299) = 8.250$ ,  $p = .004$ ). Pairwise comparisons revealed that providing choice significantly increased the perception of fluency only

when there were conversational breakdowns ( $M_{choice} = 4.17$  vs.  $M_{no-choice} = 3.34$ ,  $F(1,299) = 21.656$ ,  $p < .001$ ). Without conversational breakdowns, providing choices did not have any impact on perceived fluency ( $M_{choice} = 6.04$  vs.  $M_{no-choice} = 5.94$ ,  $F(1,299) = .383$ ,  $p = .5$ ). These results were in line with Hypothesis 2.

Next, we formally tested the full moderated mediation model using a custom model from PROCESS macro with a bootstrapping approach (Hayes 2013). The model included the presence of choice as an independent variable, perceived fluency as the first-level mediator, perceived competence as the second-level mediator, the two service evaluation variables as the dependent variables, and the presence of conversational breakdown as a moderator that moderates the effect of choice on perceived fluency. The analyses showed that, when there were conversational breakdowns, the presence of choice significantly increased the perception of fluency, which led to a higher perception of the chatbot's competence, thus enhancing the perception of service quality (indirect effect = .39; 95% CI = [.20, .61]) and satisfaction (indirect effect = .30; 95% CI = [.15, .48]). However, without conversational breakdowns, the indirect effects disappeared for both service quality (indirect effect = .05; 95% CI = [-.09, .21]) and satisfaction (indirect effect = .04; 95% CI = [-.07, .17]). The index of moderated mediation further supported the significance of the indirect effects on both the perception of service quality (index = -.34; 95% CI = [-.60, -.10]) and satisfaction (index = -.26; 95% CI = [-.47, -.07]). Overall, these results provided concrete evidence for the moderated mediation.



**Figure 1. Interaction effect of the presence of choices and conversational breakdowns**

### 3.4 Discussion

Study 1 delved into a fluency-relevant boundary condition for the impact of providing choices: conversational breakdowns. While finding evidence for the moderating role of conversational breakdowns,

the study also revealed the role of perceived fluency as a driving force for such moderating effect, which spills over to the perception of the chatbot's competence. In sum, choices can amplify a customer's perception of fluency that might have been disrupted due to a conversational breakdown. The increased perception of fluency enhances the perception of the chatbot's competence, thus ultimately improving service evaluations, only when choices are provided after conversational breakdowns.

While we discovered a boundary condition in which providing choices has a positive impact, it is also possible that the presence of choice engenders a negative impact by hurting the perception of fluency. This can happen, for instance, when a given choice set is not comprehensive. We focus on this boundary condition in the subsequent study.

## 4. Study 2

This study aimed to examine the role of choice comprehensiveness for the impact of providing choices on service evaluations. As discussed previously, the benefit of providing choice after a conversational breakdown may backfire if the provided choice set hurts the perception of fluency, for instance, due to its incomprehensiveness. Thus, we keep the existence of conversational breakdowns in this study across conditions, while manipulating the presence of choice and the comprehensiveness of a choice set in a between-subjects design. Similar to Study 1, we also study whether the perceptions of fluency and the chatbot's competence serve as underlying mechanisms for the proposed effect.

### 4.1 Stimulus materials, procedures, and measures

One hundred and eighty-one undergraduate students (107 female) from a U.S. university participated in the study in exchange for course credit. Participants were randomly assigned to one of the three conditions: comprehensive or incomprehensive choice sets, and no-choice as a control.

While we used a similar predesigned script from Study 1, we used a different cover story to extend the generalizability and efficiently create a variation in choice comprehensiveness. The scenario described a recent order of a textbook, which needed to be exchanged for a newer edition. In addition, it also required the participants to request a free shipping label for returning the product to be exchanged. This additional task was added to manipulate choice comprehensiveness more frequently. Furthermore, in

practice, a choice set is likely to be incomprehensive as customers' service requests become more complex, because a predefined choice set is unlikely to encompass every potential request.

To manipulate choice comprehensiveness, we varied the content of available choices by replacing the option the participants were supposed to choose to the option that are realistic but does not correspond to the issue described in the cover story. For instance, when the chatbot asked if there is anything else needed after processing the exchange request, participants had to type in their message. Then, a breakdown occurs, and the chatbot provides three options: "express shipping for the delivery (\$15)", "subscribe to the newsletter for discounts and offers", and "free shipping label for the return" for the comprehensive choice condition. For the incomprehensive choice condition, the last option was replaced with "create a membership". Other than the content of one of the options, we kept all other elements of choice architecture constant (Johnson et al. 2012). The choice was provided only after a conversational breakdown as our goal was to examine whether the effectiveness of choice discovered in Study 1 persists regardless of the choice set. Those in the no-choice condition encountered the same messages from the chatbot, except the provision of choice. Instead, the participants in the no-choice conditions were asked to freely type in their response.

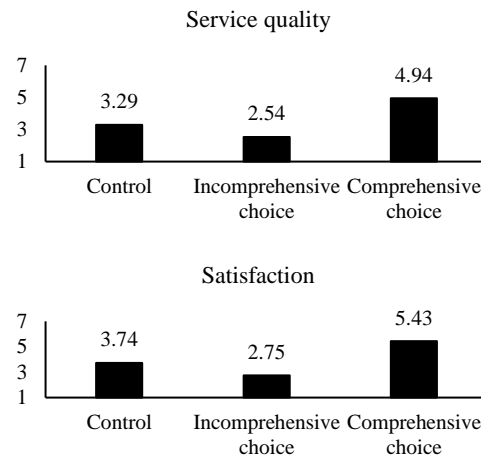
Along with the measures used in the prior study, we measured participants' perception of choice comprehensiveness using three items (e.g., "sufficient for completing the task"; "Strongly disagree / Strongly agree") on a seven-point semantic differential scale as manipulation check (Yang et al. 2005). The perception of choice comprehensiveness was measured only for those who were assigned to the two choice conditions.

## 4.2 Results

One hundred and fifty subjects passed the two attention checks and were used in the analyses. We first confirmed our manipulations by finding that participants in the choice conditions perceived that they encountered choices more frequently than those in the control condition ( $M_{choice} = 2.89$  vs.  $M_{no-choice} = 1.29$ ,  $SDs = .85$  and  $.64$ ,  $t(148) = 11.930$ ,  $p < .001$ ) and that those in the incomprehensive choice condition perceived the choices provided to be less comprehensive than those in the comprehensive choice condition ( $M_{incomprehensive} = 2.52$  vs.  $M_{comprehensive} = 5.91$ ,  $SDs = 1.30$  and  $.90$ ,  $t(96) = -15.202$ ,  $p < .001$ ).

To test our third hypothesis about the effect of choice comprehensiveness, we conducted pairwise comparisons across conditions with two service evaluation variables as the dependent variables. First,

we observed that those who encountered comprehensive choice sets after conversational breakdowns compared to those who did not encounter any choices reported greater perception of service quality ( $M_{control} = 3.29$  vs.  $M_{comprehensive} = 4.94$ ,  $SDs = 1.39$  and  $1.40$ ,  $t(103) = 6.070$ ,  $p < .001$ ) and satisfaction ( $M_{control} = 3.74$  vs.  $M_{comprehensive} = 5.43$ ,  $SDs = 1.60$  and  $1.38$ ,  $t(103) = 5.798$ ,  $p < .001$ ). These findings replicate the positive effect of choice provision after conversational breakdowns observed in Study 1. Meanwhile, we found that providing incomprehensive choice sets compared to providing no choice led to lower evaluations of service quality ( $M_{incomprehensive} = 2.54$  vs.  $M_{control} = 3.29$ ,  $SDs = 1.22$  and  $1.39$ ,  $t(95) = -2.820$ ,  $p = .006$ ) and satisfaction ( $M_{incomprehensive} = 2.75$  vs.  $M_{control} = 3.74$ ,  $SDs = 1.34$  and  $1.60$ ,  $t(95) = -3.286$ ,  $p = .001$ ). These results confirm Hypothesis 3 by revealing a backfiring impact of a chatbot's provision of an incomprehensive choice set. Figure 2 illustrates the results.



**Figure 2. The effect of choice comprehensiveness in Study 2**

Next, to test the underlying mechanism, we used a serial mediation model (Model 6) from PROCESS macro with a bootstrapping approach (Hayes 2013). The model included the treatment conditions as our independent variable, perception of fluency as the first-level mediator, and perception of competence as the second-level mediator. The independent variable was coded as a categorical variable with the control condition as the reference point. The analyses revealed that providing an incomprehensive choice set significantly decreased the perception of fluency, which led to lower perception of the chatbot's competence, thus hurting service quality (Indirect effect =  $-.22$ ; 95% CI =  $[-.43, -.04]$ ) and satisfaction (Indirect effect =  $-.18$ ; 95% CI =  $[-.40, -.04]$ ). On the other hand, providing a comprehensive choice set significantly improved the perception of fluency,

which led to higher perception of the chatbot's competence, thus enhancing service quality (Indirect effect = .36; 95% CI = [.16, .59]) and satisfaction (Indirect effect = .30; 95% CI = [.12, .57]). These findings altogether confirm Hypothesis 4.

### 4.3 Discussion

In Study 2, we examined another fluency-relevant boundary condition for the impact of providing choices: choice comprehensiveness. The findings supported our hypotheses by illuminating the negative impact of providing an incomprehensive choice set after conversational breakdowns on service evaluations, and such a negative impact is due to reduced perception of fluency, followed by reduced perception of the chatbot's competence. Indeed, because chatbots are programmed to provide a predefined set of choices, it is very likely that they cannot flexibly incorporate more complex customer requests. Facing a choice set that does not include the desired request, customers may feel disrupted, and the service process to be disfluent, which can lead them to believe that the chatbot is incompetent. This study, along with Study 1, underscores the role of fluency during a service interaction and the importance of choice comprehensiveness for realizing the benefit of choice provision after conversational breakdowns.

## 5. General discussion

This research investigated the role of a chatbot's choice provision during a service interaction. Based on the notion of fluency (Alter and Oppenheimer 2009), we proposed two boundary conditions—conversational breakdowns and choice comprehensiveness—for the impact of a chatbot's choice provision on service evaluations. We hypothesized that the two boundary conditions moderate the impact of a chatbot's choice provision by altering customers' perception of the fluency of a service experience and thus, the perception of the chatbot's competence. We proposed that providing choices can enhance service outcomes only when conversational breakdowns occur. We further argued that providing choices after conversational breakdowns may backfire when a choice set is incomprehensive. We conducted two experimental studies and found support for these hypotheses.

### 5.1 Theoretical implications

Although prevalent in practice, the implementation of choices on a service chatbot has not

been studied extensively. Prior studies on a service chatbot have focused mostly on the impact of incorporating social factors, such as anthropomorphism and human-like conversational behaviors (Crollic et al. 2022; Kim et al. 2019; Schanke et al. 2021). In fact, choice provision is a conversational characteristic unique to a chatbot and may affect customers' perception of a chatbot and its performance. Implementing choice has also been suggested as one of the repair strategies after chatbot failures (Ashktorab et al. 2019; Benner et al. 2021), but we are not aware of any empirical efforts investigating its effects. By exploring the impact of choice implementation and its boundary conditions, we provide a more complete picture of when and why implementing choices can be a boon or a bane. We also apply comprehensiveness, which is primarily studied in organizational decision-making contexts (Forbes 2007), to the context of individual decision-making and suggest choice comprehensiveness as an important element of choice architecture. More importantly, we question the conventional wisdom that implementing choice is always better, thus extending the understanding of a prevalent characteristic of a service chatbot and its impact on users' assessment of a chatbot's performance.

Our research also contributes to customer service literature, specifically to the stream about technology-induced service interactions, which has been mostly confined to e-service or self-service technology (Bitner et al. 2000; Meuter et al. 2000). Because of the uncertainty of a service environment, the current technology cannot perfectly avoid potential failures, and thus, conversational breakdowns are inevitable (Ashktorab et al. 2019). Also, because chatbots and a predefined set of choices are not flexible enough to satisfy a complex request, always providing a comprehensive list of choices is challenging. Thus, it is crucial to understand how the implementation of choice interacts with these boundary conditions. In addition to revealing the interactions, we illuminated that the impacts of these moderations on service outcomes occur because of the perception on the fluency of a service delivery process, which spills over to the perception of the chatbot's competence. These findings altogether add to the nascent literature on failures during a service encounter with AIs (Choi et al. 2020; Sheehan et al. 2020).

Broadly, our research bolsters the literature on fluency. While cues for fluency can take various forms, from visual to linguistic to semantic (Alter and Oppenheimer 2009), we present novel cues for fluency unique to the context of chatbot-initiated service interaction. Through the findings related to the interaction of choice implementation and the two



contextual variables, we identify boundary conditions that can either augment or deteriorate people's perception of fluency. We also reveal how the perception of fluency can be associated with the competence of an agent which engages in an act that drives fluency. We further highlight the impact of fluency on people's judgments of their service experience by showing the mediated moderating role of fluency on service evaluations.

## 5.2 Practical implications

Our work presents valuable guidance for practitioners who have deployed or are considering deploying service chatbots. While intuition suggests that choice implementation streamlines a service delivery process and helps deliver a satisfactory service experience, it does not necessarily provide any value for customers during an 'ideal' service interaction. Its value is only recognized and achieved when the perception of fluency is impaired (e.g., conversational breakdowns). Choice implementation can, in fact, be treated as a recovery strategy to reduce the potential negative consequences of chatbot failures in general. From a firm's perspective, providing choice not only alleviates the negative impact of conversational breakdowns but also leads to a more structured conversation and prevents further failures. Thus, unlike other recovery strategies from the prior literature, such as apologizing or providing explanations (Choi et al. 2020), choice provision can be a cost-efficient solution that provides a higher chance of getting the conversation back to track and complete the service process satisfactorily.

On the other hand, our findings alarm practitioners by highlighting how a chatbot's choice provision can be counterproductive if the provided choice set is not comprehensive. Providing choice has a limitation because a chatbot cannot accurately anticipate every customer request and has to rely on a predefined, limited set of choices. Due to the likelihood of an incomprehensive choice set, choice provision can backfire when a customer wants to make a complicated or unique request, thus disrupting service delivery. This implies that firms should not haphazardly implement choices to realize the potential value of mitigating the negative consequences of conversational breakdowns. Overall, we urge firms to carefully assess the right timing and occasion for providing choice to enhance the benefits while weighing its potential costs.

## 6. References

- Addas, S., and Pinsonneault, A. 2015. "The Many Faces of Information Technology Interruptions: A Taxonomy and Preliminary Investigation of Their Performance Effects," *Information Systems Journal* (25:3), pp. 231-273.
- Alter, A. L., and Oppenheimer, D. M. 2009. "Uniting the Tribes of Fluency to Form a Metacognitive Nation," *Personality and Social Psychology Review* (13:3), pp. 219-235.
- Ashktorab, Z., Jain, M., Liao, Q. V., and Weisz, J. D. 2019. "Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns," in: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, Scotland UK: Association for Computing Machinery, p. Paper 254.
- Benner, D., Elshan, E., Schöbel, S., and Janson, A. 2021. "What Do You Mean? A Review on Recovery Strategies to Overcome Conversational Breakdowns of Conversational Agents," *ICIS 2021*.
- Bitner, M. J., Brown, S. W., and Meuter, M. L. 2000. "Technology Infusion in Service Encounters," *Journal of the Academy of Marketing Science* (28:1), pp. 138-149.
- Brady, M. K., and Cronin, J. J. 2001. "Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach," *Journal of Marketing* (65:3), pp. 34-49.
- Chambon, V., and Haggard, P. 2012. "Sense of Control Depends on Fluency of Action Selection, Not Motor Performance," *Cognition* (125:3), pp. 441-451.
- Choi, S., Mattila, A. S., and Bolton, L. E. 2020. "To Err Is Human(-Oid): How Do Consumers React to Robot Service Failure and Recovery?," *Journal of Service Research*, p. 1094670520978798.
- Crolic, C., Thomaz, F., Hadi, R., and Stephen, A. T. 2022. "Blame the Bot: Anthropomorphism and Anger in Customer-Chatbot Interactions," *Journal of Marketing* (86:1), pp. 132-148.
- Fernández-Sabiote, E., and López-López, I. 2020. "Discovering Call Interaction Fluency: A Way to Improve Experiences with Call Centres," *Service Science* (12:1), pp. 26-42.
- Fiske, S. T., Cuddy, A. J. C., and Glick, P. 2007. "Universal Dimensions of Social Cognition: Warmth and Competence," *Trends in Cognitive Sciences* (11:2), pp. 77-83.
- Forbes, D. P. 2007. "Reconsidering the Strategic Implications of Decision Comprehensiveness," *Academy of Management Review* (32:2), pp. 361-376.
- Fotheringham, D., and Wiles, M. A. 2022. "The Effect of Implementing Chatbot Customer Service on Stock Returns: An Event Study Analysis," *Journal of the Academy of Marketing Science*.
- Fredrickson, J. W. 1984. "The Comprehensiveness of Strategic Decision Processes: Extension, Observations, Future Directions," *Academy of Management Journal* (27:3), pp. 445-466.
- Froehle, C. M., and White, D. L. 2014. "Interruption and Forgetting in Knowledge-Intensive Service

- Environments," *Production and Operations Management* (23:4), pp. 704-722.
- Graf, L. K. M., Mayer, S., and Landwehr, J. R. 2018. "Measuring Processing Fluency: One Versus Five Items," *Journal of Consumer Psychology* (28:3), pp. 393-411.
- Han, E., Yin, D., and Zhang, H. Forthcoming. "Bots with Feelings: Should Ai Agents Express Positive Emotion in Customer Service?," *Information Systems Research*.
- Hayes, A. F. 2013. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY, US: Guilford Press.
- Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, W. E., and Schlesinger, L. A. 1994. "Putting the Service Profit Chain to Work," in: *Harvard Business Review*. pp. 164-174.
- Huang, M.-H., and Rust, R. T. 2021. "Engaged to a Robot? The Role of Ai in Service," *Journal of Service Research* (24:1), pp. 30-41.
- Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C., Goldstein, D. G., Häubl, G., Larrick, R. P., Payne, J. W., Peters, E., Schkade, D., Wansink, B., and Weber, E. U. 2012. "Beyond Nudges: Tools of a Choice Architecture," *Marketing Letters* (23:2), pp. 487-504.
- Kim, S. Y., Schmitt, B. H., and Thalmann, N. M. 2019. "Eliza in the Uncanny Valley: Anthropomorphizing Consumer Robots Increases Their Perceived Warmth but Decreases Liking," *Marketing Letters* (30:1), pp. 1-12.
- Klopfenstein, L. C., Delpriori, S., Malatini, S., and Bogliolo, A. 2017. "The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms," in: *Proceedings of the 2017 Conference on Designing Interactive Systems*. Edinburgh, United Kingdom: Association for Computing Machinery, pp. 555-565.
- Kulkarni, A. 2023. "The Good, the Bad and the Bot: How Chatbot Experiences Can Make or Break Your Business," in: *Forbes*.
- Luo, X., Tong, S., Fang, Z., and Qu, Z. 2019. "Frontiers: Machines Vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases," *Marketing Science* (38:6), pp. 937-947.
- Markets and Markets. 2019. "Chatbot Market by Component, Usage, Technology, Deployment Model, Application, Organization Size, Vertical, and Region - Global Forecast to 2024."
- Meuter, M. L., Ostrom, A. L., Roundtree, R. I., and Bitner, M. J. 2000. "Self-Service Technologies: Understanding Customer Satisfaction with Technology-Based Service Encounters," *Journal of Marketing* (64:3), pp. 50-64.
- Mordor Intelligence. 2022. "Chatbot Market-Growth, Trends, Covid-19 Impact, and Forecasts (2022-2027)."
- Nguyen, Q. N., Sidorova, A., and Torres, R. 2022. "User Interactions with Chatbot Interfaces Vs. Menu-Based Interfaces: An Empirical Study," *Computers in Human Behavior* (128), p. 107093.
- Oliva, T. A., Oliver, R. L., and Macmillan, I. C. 1992. "A Catastrophe Model for Developing Service Satisfaction Strategies," *Journal of Marketing* (56:3), pp. 83-95.
- Oppenheimer, D. M. 2008. "The Secret Life of Fluency," *Trends in Cognitive Sciences* (12:6), pp. 237-241.
- Ostrom, A. L., Field, J. M., Fotheringham, D., Subramony, M., Gustafsson, A., Lemon, K. N., Huang, M.-H., and Mccoll-Kennedy, J. R. 2021. "Service Research Priorities: Managing and Delivering Service in Turbulent Times," *Journal of Service Research* (24:3), pp. 329-353.
- Parasuraman, A., Zeithaml, V. A., and Berry, L. L. 1985. "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing* (49:4), pp. 41-50.
- Schanke, S., Burtch, G., and Ray, G. 2021. "Estimating the Impact of "Humanizing" Customer Service Chatbots," *Information Systems Research* (32:3), pp. 736-751.
- Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., and Simons, A. 1991. "Ease of Retrieval as Information: Another Look at the Availability Heuristic," *Journal of Personality and Social Psychology* (61:2), pp. 195-202.
- Seshadri, S., and Shapira, Z. 2001. "Managerial Allocation of Time and Effort: The Effects of Interruptions," *Management Science* (47:5), pp. 647-662.
- Sheehan, B., Jin, H. S., and Gottlieb, U. 2020. "Customer Service Chatbots: Anthropomorphism and Adoption," *Journal of Business Research* (115), pp. 14-24.
- Shen, X.-L., Li, Y.-J., Sun, Y., and Wang, N. 2018. "Channel Integration Quality, Perceived Fluency and Omnichannel Service Usage: The Moderating Roles of Internal and External Usage Experience," *Decision Support Systems* (109), pp. 61-73.
- Simonite, T. 2017. "Facebook's Perfect, Impossible Chatbot," in: *MIT Technology Review*.
- Speier, C., Valacich, J. S., and Vessey, I. 1999. "The Influence of Task Interruption on Individual Decision Making: An Information Overload Perspective," *Decision Sciences* (30:2), pp. 337-360.
- Stewart, D. M., and Chase, R. B. 1999. "The Impact of Human Error on Delivering Service Quality," *Production and Operations Management* (8:3), pp. 240-263.
- Thaler, R. H., and Sunstein, C. R. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press.
- Winkielman, P., Schwarz, N., Fazendeiro, T. A., and Reber, R. 2003. "The Hedonic Marking of Processing Fluency: Implications for Evaluative Judgment," in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, J.M.K.C. Klauer (ed.). Mahwah, NJ: Lawrence Erlbaum Associates, pp. 189-217.
- Xu, J., Benbasat, I., and Cenfetelli, R. T. 2014. "Research Note—the Influences of Online Service Technologies and Task Complexity on Efficiency and Personalization," *Information Systems Research* (25:2), pp. 420-436.
- Yang, Z., Cai, S., Zhou, Z., and Zhou, N. 2005. "Development and Validation of an Instrument to Measure User Perceived Service Quality of Information Presenting Web Portals," *Information & Management* (42:4), pp. 575-589.