

Master Thesis

Unraveling the Impact: Customer Confusion and the Shopping Bot Experience

A Wizard-of-Oz Experiment Using Eye Tracking



Radboud Universiteit

Name: Esther van Gesink

Student number: S1044398

Supervisor: Dr. K. Sidaoui

Second examiner: R.W.H. Wetzels MSc

Radboud University, Faculty of Management

MSc Marketing

Nijmegen, 30/06/2023

Abstract

This thesis examines the role of customer experience (CX) in the relationship between customer confusion on satisfaction and loyalty, in a self-service technology setting. Currently, an increase in the use of chatbots by companies causes customer confusion to happen more often during shopping experiences. Consequently, confusing chatbot conversations cause dissatisfaction and disloyalty towards the company. Therefore, this thesis demonstrated that the negative outcomes caused by confusion can be partially explained by a customer's experience, by conducting a lab experiment. During the experiment, the participant went through a shopping experience with a chatbot. Customer confusion and time pressure were manipulated to create four conditions. The findings show a strong mediating relationship of thoughts on the relationship between customer confusion and satisfaction and loyalty. No relation was found for the eye sensations of the customer. Furthermore, the results revealed that the age of the customer has a significant effect on their thoughts, making older people more prone to evaluate the service more negatively. This research contributes to the body of knowledge by showing the mediating effect of thoughts on the already established relationship between customer confusion and satisfaction and loyalty. Additionally, it provides evidence that firms should take into account their customer's experience when implementing a chatbot since confusion can do damage to a firm financially.

Keywords: customer confusion, customer experience, self-service technology, eye tracking

Contents

1. Introduction	5
1.1 Research Objective.....	7
1.2 Relevance	8
1.3 Outline.....	10
2. Literature Review	10
2.1 Chatbots as Self-Service Technology	10
2.2 Customer Confusion from Chatbots.....	12
2.3 Customer Experience	14
2.3.1 Customer Satisfaction	14
2.3.2 Customer Loyalty.....	15
2.4 Customer Experience Phenomenon	16
2.4.1 Customer Confusion and Sensations.....	17
2.4.2 Thoughts.....	19
2.5 Customer Confusion and its Influence on Outcomes via Customer Experience	20
2.6 Time Pressure	22
2.7 Conceptual Model	22
2.7.1 Control Variables	23
3. Methodology	24
3.1 Procedure.....	25
3.2 Participants	26
3.3 Design.....	27
3.4 Measurement	28
3.5 Data Analysis Procedure	28
3.6 Research Ethics and Limitations	29
4. Results	30
4.1 Measurement Model.....	30
4.1.1 Control Variables	33
4.2 Structural Model.....	34
4.2.1 Mediation Analysis	36
4.2.2 Multi-group Analysis with Time Pressure	36
5. Discussion	37

5.1 Theoretical Implications.....	37
5.2 Managerial Implications.....	39
5.3 Limitations	41
5.4 Future Research.....	42
5.5 Conclusion.....	44
References	45
Appendix A. Manipulation.....	61
Table A1. <i>Independent sample t-test outcome</i>	61
Figure A1. <i>Condition one (no confusion and no time pressure)</i>	62
Figure A2. <i>Condition two (yes confusion and no time pressure)</i>	63
Figure A3. <i>Condition three (no confusion and yes time pressure)</i>	63
Figure A4. <i>Condition four (yes confusion and yes time pressure)</i>	64
Appendix B. Operationalization Table.....	65
Appendix C. Consent Form.....	70
Appendix D. Script Chatbot	71
Appendix E. Data Preparation.....	76
Table E1. <i>Univariate statistics</i>	76
Appendix F. Results	77
Table F1. <i>Fornell-Larcker criterion</i>	77
Table F2. <i>Heterotrait-monotrait ratio (HTMT)</i>	77
Table F3. <i>Path coefficients and significance</i>	77
Table F4. <i>Path coefficient, t-value, and p-value of control variables</i>	78

1. Introduction

Since the COVID-19 pandemic, companies needed to be creative to reduce physical interactions with their customers, resulting in rapid growth in the digitalization of companies (Verma & Gustafsson, 2020). Digital technologies, such as self-service technologies (SST), were and are used incrementally by companies to adjust to the changes in the world. Besides, these recent technologies have changed the service market as well (Almeida et al., 2020). SSTs make it possible for customers to experience a service without employee interaction while still improving customer experience, which was a necessity during the pandemic (Meuter et al., 2005). The customer experience (CX) refers to the different responses someone has to a service, which can be cognitive, affective, physical, sensorial, and social (Becker & Jaakkola, 2020; Lemon & Verhoef, 2016). SSTs are a fantastic way to maintain a positive CX, even without the physical touchpoints with the employees. Moreover, they reduce important resources of the company, such as time and money (Beatson et al., 2007). Consequently, after the pandemic, these technologies, like chatbots, were still used and seen as beneficial by businesses. With an annual growth of around 23 percent (Durach & Diachenko, 2023), by 2027, chatbots will become the primary customer service channel for about 25 percent of all organizations (Gartner, 2022). Customers have embraced these technologies themselves as well, as they can bring convenience, control, and privacy (Dabholkar & Spaid, 2012).

So, chatbots have changed the way customers experience services for the good, but it also has their downsides (Beatson et al., 2007). As customers are often not familiar with these technologies and have to process this information on their own, this can lead to confusion (Mitchell & Papavassiliou, 1999). Identified as the failure of the customer to correctly interpret various aspects of a good or service during the information-processing process (Turnbull et al.,

2000). Since customer confusion is an effect of service failure, post-purchase behavior, like customer satisfaction and loyalty, can be impacted (Foxman et al., 1992; Matzler & Waiguny, 2005; Smith & Bolton, 1998). Therefore, confusion is of huge interest to marketing managers, making it an essential problem to tackle (Walsh & Mitchell, 2010). Moreover, it is expected that customer confusion affects the CX, which in turn influences the satisfaction and loyalty of the customer (Anninou & Foxall, 2019; Becker & Jaakkola 2020; Innis & La Londe, 1994; Lemon & Verhoef, 2016).

Since CX is a broad concept, a focus will be on sensations and thoughts as these concepts reflect upon each other, but still expected to occur sequentially (Meißner & Oll, 2017). Sensational responses, such as eye movements, have proven to reflect how confused customers are (Meißner & Oll, 2017; Rozin & Cohen, 2003), affecting the person's eye movements and visual attention (Pachman et al., 2016). This gaze behavior occurs during the experience, so in real-time (Lemon & Verhoef, 2016). Cognition, however, does occur thereupon. It is the result of processing the interaction with the firm into thoughts, ideas, memories, language, and problem-solving (Bustamante & Rubio, 2017; Keiningham et al., 2017). Since cognition covers several processes, thoughts will be focused upon, as confusion is expected to reflect a lack of clear and orderly thought as well (Leek & Kun, 2006). Furthermore, existing literature on confusion focuses on affective responses (Anninou & Foxall, 2019). Consequently, there is a lack of research measuring the effect of customer confusion on cognition and thought in particular (Varela-Neira et al., 2008).

Additionally, recent studies on confusion used methods such as interviews and surveys, as a consequence there is a need for research combining scenario-based experiments with real behaviors (Grégoire & Matilla, 2021). This study attempts to measure how sensations and the thought process of the customer are affected by confusion. And how this will result in

dissatisfaction and a decrease in loyalty, by using a combination of subjective data of thoughts by scenario-based research, with the objective data conducted by measuring real behavior, such as thoughts and sensations (Grégoire & Mattila, 2021).

1.1 Research Objective

How CX affects the relationship between customer confusion and business outcomes remains unclear. Firstly, the literature lacks research on the role of CX in the relationship between customer confusion and customer satisfaction and loyalty. Especially, the effect of customer confusion on CX has not been studied in detail (Varela-Neira et al., 2008). Studying this gap will help identify why confusion arises and how it leads to dissatisfied and disloyal customers, which is important information for marketing managers to help improve the company (Walsh & Mitchell, 2010). Secondly, there is a gap concerning the context of the studies about customer confusion. Often, chatbots are not taken into consideration when studying customer confusion, however, it could be beneficial for companies since these technologies are used incrementally and could be experienced differently than traditional services (Cunningham et al., 2009). Lastly, since most studies on customer confusion are conducted through interviews or surveys, research is needed based on experiments and real behaviors (Grégoire & Matilla, 2021). The use of novel technologies, like eye tracking and chatbots, can give new insights into how people experience confusion from a CX perspective. These insights could help firms to understand where their negative business outcomes come from and adjust their chatbots to prevent these outcomes from happening.

Having identified these gaps, the present study considers CX while exploring the relationship between customer confusion and customer satisfaction and loyalty. A focus on CX will provide a clearer understanding of customer confusion in a shopping experience context.

Moreover, identifying how sensations and thoughts cause confusion to escalate into customer dissatisfaction and disloyalty with the firm, gives an understanding of the use of chatbots. Since the focus is on the shopping experience, a shopping bot will be used, which is a supportive chatbot helping with locating and comparing products (Rowley, 2000). Using chatbots will help to understand the CX as it is seen as a reliable tool to measure and analyze customers' responses (Sidaoui et al., 2020). Hence, it helps firms understand how their services affect their customers, by examining the CX, and gives insights into how the technology can be used to recover from customer confusion. To conduct this study, the following research question will be used: *How does customer confusion impact the customer experience with a shopping bot?*

This study aims to answer this question by experimenting with a shopping bot in a situation that stimulates confusion. During this experiment eye tracking is used to measure gaze behavior and chatbots help to communicate with the individual to get to know their thought process during the experiment.

1.2 Relevance

This study contributes to the body of knowledge on customer confusion and CX in several ways. Firstly, current literature lacks the use of SSTs when studying the effects of customer confusion (Cunningham et al., 2009). This study broadens the literature on customer confusion by expanding its boundaries with a chatbot context. Secondly, a negative relationship between customer confusion and CX is already established in the literature (Anninou & Foxall, 2019; Becker & Jaakkola 2020), additionally, confusion has a negative effect on post-purchase behavior (Matzler & Waiguny, 2005). Establishing an effect between customer confusion and CX (outcomes) will help explain this relationship better. Studying this effect will provide a coherent story about a customer's journey when experiencing chatbots and how it can lead to

dissatisfaction and disloyalty. Thirdly, the results of this study expand the knowledge of thoughts as part of the CX. Therefore, this study highlights the importance of CX and how its different elements can be used to research what effects new technologies can have on the customer, and the business outcomes as a result.

The results also show managerial relevance. Firstly, since chatbots have a prominent role in the service landscape, expected to be worth \$9.4 billion by 2024 (Omar, 2023), the contribution of this study will affect most businesses using novel technologies. Currently, managers are facing the challenge of dealing with the consequences of chatbots about how customers assess the company (Beatson et al., 2007). This is important to find out, as service failures, such as that of confusion, indirectly damage market share and profitability (Keaveney, 1995). Furthermore, it shows the importance of a human-centric chatbot design to enhance CX while still meeting financial needs (Blazevic & Sidaoui, 2020). Secondly, the findings provide insights into the negative consequences that can occur during self-service, its effect on CX, through sensations and thoughts, and their outcomes, like satisfaction and loyalty. Consequently, this knowledge will help prevent these negative outcomes from happening, making it possible for firms to build long-term relationships with their loyal customers and hence increase profits (Singh, 2006; Venetis & Ghauri, 2004). Lastly, this study highlights the possibility of using chatbots to improve CX by gathering useful feedback during interaction with the customer. Feedback about the thoughts of customers can improve the understanding of how the confusion came to be and thus how it can be recovered. This study is interesting for marketing managers who struggle with losing customers.

1.3 Outline

The remainder of this study will have the following structure. First, the existing literature on the concepts related to customer confusion and CX will be analyzed and discussed. By linking the concepts, relations will be identified and visualized in a conceptual model. Second, the methodology of this study will be elaborated on. Third, the results of the experiment will be reviewed and discussed. Last, theoretical and managerial implications will be listed, limitations and future research will be suggested and an answer will be given in the conclusion.

2. Literature Review

In this second chapter, the most important concepts will be discussed and linked in a conceptual model. First, the concept of SST and chatbots will be discussed. Second, the relationship between customer confusion and CX outcomes will be elaborated on. Third, CX will be discussed, concerning the confusion, satisfaction, and loyalty of customers, in addition to the effect of time pressure. Ultimately, control variables are introduced and a conceptual model is visualized based on this literature.

2.1 Chatbots as Self-Service Technology

In marketing, artificial intelligence (AI) is a rising concept. New technologies are used more often to help marketers track real-time data to analyze habits, likings, and purchases and respond to customer requirements (Verma et al., 2021). AI technology in combination with the huge amount of data coming from customers provides new opportunities for firms, like SSTs (Sidaoui et al., 2020). It has transformed interpersonal interactions with technology (Curran et al., 2003), causing an important shift in services (Parasuraman, 1996), as employees are not needed to perform the service anymore (Meuter et al., 2005). Numerous studies explain the benefits of using SST, the most prominent reason for firms is to save time and money (Beatson et al., 2007),

and for customers, place-and-time convenience (Yang & Klassen, 2008), control, and privacy (Dabholkar & Spaid, 2011).

A form of SSTs is chatbots. Hence, research on SSTs can be applied to chatbots as well. A chatbot is a computer program using AI to communicate with customers in natural language, using the learning capabilities given. The human-like characteristics derived from AI make them an interesting tool in online service encounters (Feine et al., 2019; Sidaoui et al., 2020; Verma et al. 2011). Consequently, chatbots can understand the CX and can respond appropriately (Kaczorowska-Spychalska, 2019). Hence, they replace employees and help firms save time and up to 30% in customer support costs (Shukairy, 2022).

On the contrary, negative consequences can occur from the use of chatbots. Much research has noted service failures and their big impact on customer satisfaction (e.g. Hess et al., 2003; McCollough et al., 2000; Smith & Bolton, 1998). Service failures occur when a customer is dissatisfied with a service, as they have a problem with the service since it does not meet their expectations (Bell & Zemke, 1987; Spreng et al., 1995). As customers are often unfamiliar with the technology and thus can have difficulty processing the information, confusion can occur (Mitchell & Papavassiliou, 1999). Customer confusion can be seen as a service failure, whereas it interrupts the service in a way that cannot be finished correctly.

So, research confirms that chatbots can be an antecedent for customer confusion, which in turn can lead to dissatisfaction and disloyalty of the customer. Firms need to improve their chatbots to prevent negative CX from happening (Matzler & Waiguny, 2005). Considering the ability of chatbots to extract and analyze useful data from customers about their CX (Sidaoui et al., 2020), it is an important concept for the customer confusion literature. This makes it possible for firms to retrieve information from the customer during and after the service, facilitating the

ability to recognize customer confusion when it occurs and extract information about their CX (Feine et al., 2019).

2.2 Customer Confusion from Chatbots

Although chatbots have a prominent place in the service industry (Almeida et al., 2020), often a human-centric design is lacking when a company has a financial focus. This hinders the CX since the individual's abilities are not taken into account (Blazevic & Sidaoui, 2022; Mannheim et al., 2019). Additionally, it can confuse and is seen as a common service failure associated with chatbots (Collier et al., 2015; Mitchell & Papavassiliou, 1999). To prevent this confusion, the customer should be taken into account when implementing the chatbot (Blazevic & Sidaoui, 2020).

Customer confusion occurs when a customer fails to correctly perceive numerous aspects of the service during the processing of the information (Turnbull et al., 2000). Three dimensions of customer confusion are recognized: similarity-, overload-, and ambiguity confusion (Edward & Sahadev, 2012; Matzler & Waiguny, 2005; Mitchell & Papavassiliou, 1999). Similarity confusion occurs when customers experience similarity between choices, products, and/or services during the decision-making process of the chatbot, causing indecisiveness, which can lead to confusion (Edward & Sahadev, 2012; Matzler & Waiguny, 2005; Mitchell & Papavassiliou, 1999). Overload confusion is the result of the confrontation with an information-rich environment and a high degree of marketing stimuli (Matzler & Waiguny, 2005). Confusion in this category is increased when having less time to process the information and when there are too many options to choose from (Mitchell & Papavassiliou, 1999). Ambiguity confusion arises when the communication in the chatbot of the firm is unclear, of low quality, and/or misleading (Edward & Sahadev, 2012; Matzler & Waiguny, 2005; Mitchell & Papavassiliou, 1999). Mitchel

and Papavassiliou (1999) add one other reason for customer confusion, which is related to alien environments. When people have experienced a new environment and the technologies associated with them, this can confuse them as they have not yet had any experience in this environment. This latter type of confusion is especially applicable to chatbots, as these technologies are new for most customers.

Customer confusion can have some negative consequences that can indirectly influence the profits of the company. The most common behavioral consequences mentioned in the literature are abandoning purchases, seeking additional information (Matzler & Waiguny, 2005), and using risk reduction strategies (Matzler & Waiguny, 2005; Turnbull et al., 2000). Turnbull and colleagues (2000) explain these consequences by the perceived risk to customers. Customers try to prevent confusion by searching for information beforehand, this is not possible for services you encounter spontaneously. Consequently, the perceived risk increases when this confusion has settled, thus customers perceive that the services can have a negative outcome. Resulting in a change of behavior, by searching for more information to process it. This behavior will lead to a purchase decision and post-purchase behavior, such as loyalty and satisfaction (Edward & Sahadev, 2012). As numerous researchers have proved the significant effects of customer confusion on the behaviors of confusion reduction strategies – for example, abandoning purchase, seeking additional information, and doing nothing (e.g. Matzler & Waiguny, 2005; Mitchell and Papavassiliou, 1999; Mitchell et al., 2005) – this study will look at the direct effect of customer confusion on the post-purchase effect on company-related consequences, in this case, customer loyalty and satisfaction.

Besides the direct effect of customer confusion on satisfaction and loyalty, confusion is expected to influence the experience of the customer as well, as it affects the sensations and cognitions of a person negatively. These CX responses in turn are presumed to affect the business

outcomes of satisfaction and loyalty too (Mitchell & Papavassiliou, 1999; Wedel & Pieters, 2017).

2.3 Customer Experience

Accordingly, customer confusion can have huge negative effects on how customers experience a chatbot. A confused customer is more likely to be dissatisfied with the service and, consequently, be less loyal to the firm in the future (Foxman et al., 1992; Smith & Bolton, 1998). Moreover, the CX, like sensations and thoughts, is proven to be an antecedent of these outcomes (Becker & Jaakkola, 2020; Lemon & Verhoef, 2016). Correspondently, customer confusion outcomes are related to CX outcomes because confusion is part of how the customer experiences the service (Mitchell & Papavassiliou, 1999). Furthermore, the positive effect of customer satisfaction on loyalty confirms that these two outcomes relate to and influence each other (Luarn & Lin, 2003). However, this does not mean that a satisfied customer is loyal as well (Bowen & Chen, 2001). Hence, both outcomes are studied.

2.3.1 Customer Satisfaction

Several studies have proven the direct negative effect of customer confusion on customer satisfaction (e.g. Edward & Sahadev, 2012; Matzler & Waiguny, 2005; Smith & Bolton, 1998). Service failures, like customer confusion, have a huge impact on switching behavior, affecting the post-purchase behavior negatively, and resulting in dissatisfaction with the service and the firm (Keaveney, 1995; Matzler & Waiguny, 2005). Anderson et al. (1994, p. 54) define satisfaction as “an overall evaluation based on the total purchase and consumption experience with a good or service over time”. Customer satisfaction can be reached by a positively perceived performance of the firm with a product or service (Innis & Londe, 1994). Thus, a positive CX is important to derive satisfaction, which will lead to higher profitability for the firm (Anderson et

al., 1994; Theron & Terblanche, 2010).

As identified before, customer confusion can be divided into three types: similarity, overload, and ambiguity confusion. These three types do have different effects on customer satisfaction, according to a study by Walsh and Mitchell (2008). Similarity confusion affects satisfaction negatively because of the similarity of information. Overload confusion leads to frustration, indecisiveness, and stress, making the decision difficult and dissatisfying (Turnbull et al., 2000; Walsh & Mitchell, 2008). Lastly, ambiguity confusion is assumed to cause a decreasing satisfaction with the service process and the firm, however, this was not supported (Walsh & Mitchell, 2008). Still, a direct negative effect of customer confusion on customer satisfaction can be found. Thus we hypothesize:

H1a. *Customer confusion negatively affects overall customer satisfaction.*

When customers are dissatisfied with the firm, this will affect the loyalty of the customer as well. Since loyalty has a positive relationship with profits, a decrease in loyalty will affect the firm financially as well. Thus, the firm needs to keep customers extremely satisfied to create loyal customers (Bowen & Chen, 2001).

2.3.2 Customer Loyalty

A loyal customer has a favorable attitude toward a firm, repurchases from this firm, and has positive word-of-mouth (Bowen & Chen, 2001). He or she values the service of a firm above that of competitors. The expectations of the customer of the service should thus be met (Griffin, 1995), which relates to how people experienced the service.

Besides the relation between the different CX outcomes, loyalty is assumed to be influenced by customer confusion directly, as shown by multiple studies on the effect of customer confusion on post-purchase behavior (Matzler & Waiguny, 2005). For example, similarity

confusion can cause disloyalty, as this type of confusion spurs the behavior of abandoning the service (Mitchell et al., 2005). Additionally, several studies acknowledge the relationship between the quality of chatbots on customer loyalty. When quality is perceived as low, which can be expressed in a poor representation of information, this can lead to confusion (Lewis & Mitchell, 1990). Thus, we can hypothesize:

H1b. *Customer confusion negatively affects overall customer loyalty.*

2.4 Customer Experience Phenomenon

Although the direct effect of customer confusion on satisfaction and loyalty is studied mostly, this study expects to find a reason behind this effect to be found in the CX literature. Since confusion can do damage to satisfaction and loyalty, it is important to recover from these incidents to diminish the damage to the firm's market share and profitability (Keaveney, 1995, McCollough et al., 2000). To do so, the CX should be understood. Additionally, how customers assess the service influences the outcomes of confusion as well, such as customer loyalty and satisfaction (Lemon & Verhoef, 2016).

Nowadays firms interact on multiple touchpoints with their customers, for example, chatbots. This has affected how customers experience the service and the interactions with the firm (Lemon & Verhoef, 2016). Besides the huge amount of research on CX, it is still a broad and fuzzy concept, as there is no coherent definition given in the literature. As there are unclarities concerning the scope, breadth, and antecedents of the concept, Becker and Jaakkola (2020) have provided us with a definition based on a systematic review of the literature: "non-deliberate, spontaneous responses and reactions to particular stimuli" (p. 637). Thus, CX is a reaction to stimuli from the firm, such as chatbots. How people experience chatbots can be positive and negative (Mitchell et al., 2005). Furthermore, as this experience is a

multidimensional construct, experience is much more complex than a positive or negative feeling. It can have the following responses on customers: sensational, affective, cognitive, physical, and social identity experiences (Lemon & Verhoef, 2016; Schmitt, 1999). These responses are important to consider, as they provide unique information about the experience and why some services evoke particular outcomes.

Nonetheless, the literature on customer confusion and CX is limited (Anninou & Foxall, 2019). Hence, this study focuses on this relation. Since CX is a broad concept, only the relation between confusion and sensations and thoughts is addressed. The sensational experience is closely related to confusion, it occurs during the service and often unintentionally, e.g. eye movements (Pachman et al., 2016). Besides, sensations expressed by the eyes of the customer show how information is processed in the mind of the customer. These thoughts are often formulated consciously after the service and can appear unclear or unstructured when confused (Leek & Kun, 2006).

2.4.1 Customer Confusion and Sensations

Customer confusion is expected to have a negative effect on CX (Anninou & Foxall, 2019). Consequently, it is assumed that it will also affect the sensations customers have during their experience with the firm. The sensory experience of a customer can be divided into sight, sound, touch, taste, and smell (Schmitt, 1999). Literature on confusion has found an effect on the visual senses of a person (Meißner & Oll, 2017; Pachman et al., 2018; Wachowiak et al., 2022). Meißner and Oll (2017) mention how confusion can have an impact on the expressions on someone's face, especially their eyes. The eyes can tell the flow of information in someone's head, consequently, it tells a lot about the confusion someone experiences (Ashby et al., 2016). Higher confusion ratings will reflect in a person's eye movements and visual attention, referred to

as gaze behavior (Pachman et al., 2016). Two of those eye movements are fixations, the duration of the gaze, and saccades, the quick movements between fixations (Meißner & Oll, 2016; Ziv, 2016). Confused customers tend to have relatively longer and more fixations, especially focusing on non-relevant information (Salminen et al., 2018).

Elaborating on visual processing, the individual processes visual stimuli in multiple steps. Firstly, the visual stimuli will enter the eye as a source of light through the first layer (the cornea). The image will be portrayed inverted on the back of the eye (the retina). It is converted into an electrical signal that is sent to the brain and is processed to be perceived and recognized by the individual (Moore & Zimsak, 2017; Tessier-Lavigne, 2000). When confused, people's visual attention can lose focus on the right stimuli, often moving to the environment (Wachowiak et al., 2022). When having selective visual attention, the right and helpful information may be filtered out. This reflects difficulties in processing information, and shorter saccades length, reflecting complexity (Pachman et al., 2016; Salminen et al., 2018). These visual processes can appear in two forms: bottom-up and top-down (Connor et al., 2004). Bottom-up processing starts from raw sensory stimuli from the environment and happens automatically (Ares et al., 2013, Connor et al., 2004). For example, when the environment does not provide the right stimuli to conduct a service, the person can get confused. Top-down processing is the opposite. The individual will use previous experience and knowledge to understand the visual stimuli, it is influenced by the cognition of the person and the intention of the activity, such as the perceived risk when starting the service (Ares et al., 2013; Connor et al., 2004, Theeuwes, 2010). Both forms are important when analyzing the visual perception of the individual (Moore & Zirnsak, 2017). So, it can be hypothesized:

H2. *Customer confusion positively affects the intensity of the gaze behavior of the customer.*

2.4.2 Thoughts

Where sensations arise during a service, thoughts occur after a service has happened. Customers will compare their expectations of the service with how the service turned out, these standards are often based on previous experiences and goal-directed behavior (Hoyer et al., 2020; Oliver, 1980; Keiningham et al., 2017; Varela-Neira et al., 2008). Consequently, the customer is satisfied or dissatisfied with the service, which in the long run influences their loyalty.

Most studies on CX mention the cognitive responses of a customer, thoughts are often not considered as an individual concept. Cognition can be referred to as processing information from the interaction with the firm and is a mental response that includes thinking, but also ideas, interpretations, language, problem-solving, and memories (Bustamante & Rubio, 2017; Keiningham et al., 2017; Ok et al., 2005). However, as this study reflects on how customer confusion, resulting from the use of chatbots, is processed, thoughts are the most relevant. So far, confusion is expected to influence how clear and structured the thoughts of the customer are (Leek & Kun, 2006). Additionally, firms can use chatbots to influence how customers think of the company. Since these stimuli can affect how customers transform the interaction into thoughts (Bustamante & Rubio, 2017).

Although sensations can reflect the cognitive process of an individual (Leek & Kun, 2006), thoughts themselves cannot be collected through observations alone, interaction is needed to let them express their thoughts. Several concepts have described how thoughts can affect the processing of a service (Hoyer et al., 2020; Varela-Neira et al., 2008). Specifically, how the customer experienced it and how it will reflect in post-purchase behaviors (Varela-Neira et al., 2008). It is important to keep these in mind when interacting with the customer and evaluating his or her experience.

The customer's total evaluation of the functional advantages and disadvantages of a good

or service is known as utilitarian value (Lee & Wu, 2017). An evaluation is made of the service on its efficiency and if it has met the customer's expectations (Babin et al., 1994). Utilitarian customers tend to perceive the service experience as a task since they want to create the highest value within the smallest budget and period (Lee & Wu, 2017). When the chatbot can be experienced effectively, without any hiccups, and with high values in return, the utilitarian value will be high (Lee & Wu, 2017). However, when customers experience confusion, this will lead to an ineffective way of using chatbots, resulting in a slower process and a lack of control. Since utilitarian customers have a desire for control and high efficiency, confusion will decrease this value (Lee & Wu, 2017; Tracy, 1998). Moreover, when customers experience high utilitarian value, this will positively affect customer satisfaction and repeat-purchase intention (Babin et al., 1994; Chiu et al., 2014; Lee & Wu, 2017). Accordingly, a higher utilitarian value positively affects customer satisfaction and loyalty.

Overall, most literature on service failure examines how customers react to the recovery from these failures in a cognitive way, not with a focus on thoughts. Also, none measures the effect of customer confusion on thoughts. This is a gap, as confusion arises from a combined failure to process the information and a failure of the firm to present a coherent service. Therefore, confusion is expected to reflect in a customer's unclear and unstructured thoughts (Leek & Kun, 2006). Thus:

H3. *Customer confusion negatively affects utilitarian value.*

2.5 Customer Confusion and its Influence on Outcomes via Customer Experience

Concluding, CX is related to customer confusion and CX outcomes. It is expected that customer confusion is indirectly affecting CX outcomes through CX (i.e. sensations and thoughts). Even though this assumption has not been tested, it could be validated through the direct linkages

between the concepts.

Besides the direct effect of customer confusion on satisfaction and loyalty (Foxman, et al., 1992; Smith & Bolton, 1998), it negatively affects CX as well (Anninou & Foxall, 2019). A confused customer is expected to have unclear and unstructured thoughts (Leek & Kun, 2006). Additionally, confusion is closely related to the sensations of an individual, affecting their gaze behavior (Pachman et al., 2016). These CX responses have proven to influence customer satisfaction and loyalty as well (Becker & Jaakkola, 2020; Lemon & Verhoef, 2016).

Consequently, we can hypothesize the following relations:

H4. *Gaze behavior negatively mediate the relationship between customer confusion and (a) customer satisfaction and (b) customer loyalty.*

H5. *Thoughts negatively mediate the relationship between customer confusion and (a) customer satisfaction and (b) customer loyalty.*

Moreover, sensations are expected to influence thoughts directly, since gaze behavior shows how information is processed in the mind (Lee & Kim, 2006). Also, sensations happen in real-time, whereas thoughts arise after the service. Since customer satisfaction concerns the evaluation of the CX, sensations are not expected to fully influence customer satisfaction directly, but also indirectly through the thoughts of the customer (Lemon & Verhoef, 2016). Several studies highlighted the relationship between eye movements and cognition (Salvucci, 1999; Thomas & Lleras, 2007). Thomas & Lleras (2007) even show the possibility that eye movements can influence thought in making the right decisions. Hence, it is hypothesized:

H6. *Gaze behavior and thoughts mediate the relationship between customer confusion and (a)customer satisfaction and (b) customer loyalty.*

2.6 Time Pressure

An additional concept related to customer confusion and often experienced during a shopping experience is that of time (Goeschl & Lohse, 2018). Time pressure can be identified as the perception of how much time is available to complete a task (Maruping et al., 2015). It depends on the environment and is not a personality trait, like time urgency (Maruping et al., 2015; Mohammed & Nadkarni, 2011). It is expected that when a person is given less time, this creates an urgency to complete or abandon the purchase (Xue et al., 2020). Moreover, these individuals will have less time to overthink their possibilities, thus will regulate less information and overload confusion (Matzler & Waiguny, 2005; Mitchell & Papavassiliou, 1999; Pieters & Warlop, 1999). Consequently, the duration of fixations will be lower, as they will scan the information, without taking the time to process the information well (Pieters & Warlop, 1999). This could be a possible reason for an increase in confusion (Goeschl & Lohse, 2018). So, it can be hypothesized:

H7. *The greater the time pressure concerning the shopping experience, the stronger the effect of customer confusion on (a) gaze behavior, (b) thoughts, (c) customer satisfaction, and (d) customer loyalty.*

2.7 Conceptual Model

This chapter has discussed the most important concepts and their linkages with each other. However, control variables are needed to limit the influence of external concepts on customer confusion and CX. Taking together the hypotheses stated above and these control variables, a conceptual model can be visualized (Figure 1).

2.7.1 Control Variables

Age. A common demographic variable used to differentiate participants in technology studies is age. Older customers can be associated with a lack of confidence, knowledge, interest, and experience concerning the use of chatbots and have difficulties learning new technology (Dean, 2008; Ellis & Allaire, 1999; Venkatesh et al., 2012). Additionally, they prefer human interaction over chatbots, in comparison with younger people who can even dislike employee contact as they grew up adapting to new technologies (Morris & Venkatesh, 2000; Simon & Usunier, 2007).

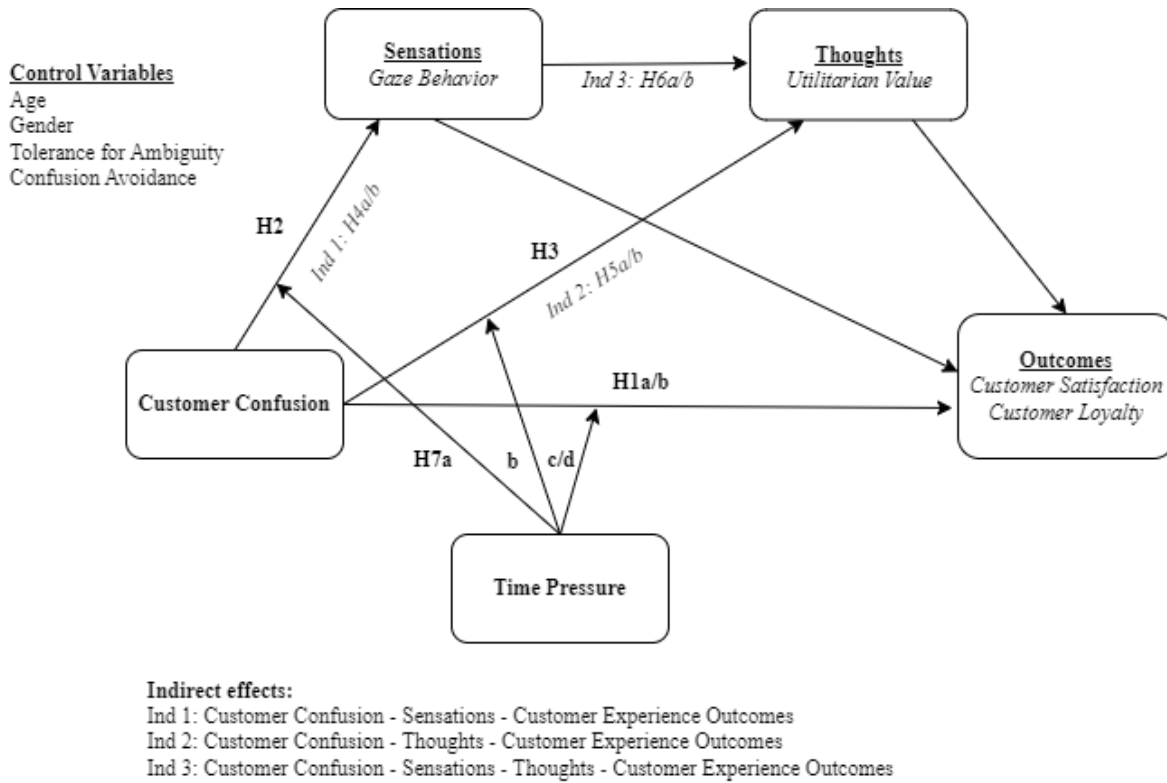
Gender. Another demographic interesting to investigate is gender. Only a few studies show how females tend to have higher levels of customer confusion when confronted with novel products or services (Turnbull et al., 2000; Wobker et al., 2015). Although others do not find a gender difference (Foxman et al., 1990). Since chatbots can be seen as unfamiliar by some, this demographic should be taken into account (Mitchell & Papavassiliou, 1999).

Tolerance for ambiguity. The degree to which someone is comfortable in a confusing, uncertain, and unpredictable situation is called tolerance for ambiguity, as clarified in the article by McLain et al. (2015). It is the ability someone has to think and behave effectively in ambiguous environments with confusing information. Ambiguous information is hard to process and hence stress can occur when not tolerant of (Furnham & Ribchester, 1995). However, customers with ambiguity tolerance will be able to think openly when confronted with confusing information and complex issues (Furnham & Ribchester, 1995).

Confusion avoidance. Confusion avoidance is the personality trait that causes people to leave a confusing or difficult circumstance (Hoque et al., 2022). People use this behavior to protect themselves from confusion, however, they cannot prevent them from having confusion (Schweizer et al., 2006).

Figure 1

Conceptual model



3. Methodology

The current study has aimed to establish a causal relationship between customer confusion and CX (outcomes), in a chatbot context. The conceptual model was tested with a quantitative approach. Experimental data was gathered by manipulating customer confusion and time pressure, in a Wizard of Oz set (Riek, 2012). Scenario-based research clarified the context in which customer confusion could occur, and eye tracking showed how confusion reflects in someone's eyes. Chatbots were used to start a conversation with the customer and collect objective data about the thoughts of the customer, as it is proven to be a reliable way of interviewing individuals about their experience (Sidaoui et al., 2020).

3.1 Procedure

Customer confusion can be expressed in someone's eyes (Meißner & Oll, 2017). Eye tracking is a useful tool to measure customer confusion in real time. Eye movements can be perceived as smooth in our own eyes, but they consist of two different components: fixations and saccades (Wedel & Pieters, 2017). Fixations are moments when the eye stays still, staring at one area, usually lasting around 100-400 milliseconds (Rayner, 1998). Saccades are fast movements of the eye (Wedel & Pieters, 2017). When confused, people have relatively longer and more fixations, especially focusing on non-relevant information, which reflects the difficulty in processing information. Consequently, there will be difficulties processing the information, because the right information is filtered out (Pachman et al., 2016; Salminen et al., 2018).

Whereas sensations can be measured from observation, other CX responses, like that of thoughts, can only be communicated by the participant. Besides the use of chatbots as a means of creating confusion, the chatbot also had the function to interview the respondents (Sidaoui et al., 2020). A Wizard of Oz technique was used for this matter. As the chatbot was remotely operated by one of the researchers (Riek, 2012). This made it able to retrieve customer information about the various variables from the conceptual model, during and after the shopping experience (Feine et al., 2019).

Upon entry to the lab, the participants were welcomed by one of the researchers and asked to fill in a consent form (see Appendix C). When permission was granted, the participant was given a participant ID that was randomly selected and asked to give this number to the chatbot when starting, as the specific number was allocated to a condition the participant would undergo. Following, the participant entered the lab and was seated. One of the researchers installed the eye tracker and calibrated it to fit the participant. After succeeding in a comfortable fit, the researcher

gave some instructions to the participant (for example, do not leave the room during the experiment, do not touch the eye tracker, etc.) and left the room. The interaction started with the chatbot giving a short introduction, followed by small talk (see Appendix D for full script). The participant was sent a link to a website to complete a purchase. After putting the purchase in their basket, an error screen occurred. In response, the chatbot was able to answer any question the participant had. Finally, the participant would fill in a number that would be accepted by the chatbot, fixing the error. Consequently, the participant was asked a few closed-end survey-type and open questions from the chatbot about the remaining constructs of thoughts, satisfaction, loyalty, and the control variables. After finishing these questions, the researcher would step in and help remove the eye tracker, debrief, and thank the participant for their time.

3.2 Participants

The data used by this experiment was obtained by conducting an experiment on the students and employees of the Radboud University in Nijmegen, the Netherlands. Non-probability convenience sampling was used since the experiment was bound to an office in one of the buildings of the Radboud University (Malhotra et al., 2006). Potential participants were approached by the researchers in the public spaces of the university. The desired level of the sample size, according to the Central Limit Theorem should be 30 per condition, to predict the characteristics of the population more accurately (Ganti, 2017; Kwak & Kim, 2017). Since this study has four conditions, a sample size of a minimum of 120 is desired, but 200 is preferred to take into account potential missing data.

The researchers were able to convince 200 respondents to participate in the study.

However, 3 respondents were eliminated due to missing data from the eye tracker, leaving a total

sample of 197 participants. The participants' average age was 25 years old and 41.6% of them were female.

3.3 Design

This lab experiment used a 2 (confusion: no vs. yes) x2 (time pressure: no vs. yes), between-subjects experimental design. Confusion was manipulated by providing multiple potential error codes on their screen. Causing information overload and unclarity since none of them was called 'error'. Time pressure was manipulated with the use of a timer, which occurred on the error page, stimulating the feeling of time pressure to solve the error. To make sure that these concepts were measured correctly, a pre-test was run before the experiment, providing significant results. Additionally, during the experiment a manipulation check was run, and analyzed by doing an independent samples t-test in IBM SPSS (Version 28), comparing the means of the groups (Ross & Wilson, 2017). For an overview of the results and the manipulation questions see Appendix A. In sum, the concepts of confusion and time pressure were manipulated correctly, as only one question about time pressure showed an insignificant difference.

Consequently, the participant was randomly assigned to one of the four conditions (see Figures A1 till A4). In the first condition, the participant experienced no confusion and no time pressure. In the second condition, confusion was evoked, but no time pressure. In the third condition, there was no confusion but there was a timer. In the last condition, an overwhelming number of incorrect numbers were shown and a timer was displayed.

The experiment took place over two weeks, during the months of May and June in 2023. It took around 20 minutes per participant. It yielded 50 valid responses for the first condition, 50 for the second condition, 49 for the third condition, and 48 for the last condition.

3.4 Measurement

To fit the setting of the chatbot context, the study measures were modified from existing literature (see Appendix B). Confusion and time pressure were manipulated during the experiment. For the sensations, an eye tracker was used, measuring the total number and the average duration of fixations of the participant on the screen (relevant area) and outside the screen (non-relevant areas; Salminen et al., 2018). The eye tracking glasses and the software Pupil Player v3.5.1 of Pupil Labs (2023) were used to measure the fixations. Thoughts were measured by open (three items) and scaled questions (four items). The open-ended questions were formulated to measure thoughts more in-depth, making it possible to provide space for the participants to describe their thoughts openly (Fossey et al., 2002; Leiva et al., 2006). The utilitarian value was measured by adapting four scale items ranging from 1 ‘strongly disagree’ to 7 ‘strongly agree’ used by Babin et al. (1994). Customer loyalty (two items) was based on work by Reichheld (2003) and satisfaction (one item) was measured by asking about the degree of satisfaction. Both outcomes relied on seven-point semantic scales. The chatbot also asked questions about age, gender, tolerance for ambiguity, confusion avoidance, and alternative explanations.

3.5 Data Analysis Procedure

Before the analysis, the gathered data was prepared with the use of IBM SPSS (Version 28; 2021). The dataset was checked for any missing data, since three participants had no eye tracking data, these were deleted beforehand. Next, a missing value analysis was run (see Appendix E). All missing data was less than 5% and missing completely at random (sig. = 0.062), hence, we do not consider it a problem (Roth & Switzer, 1995). Additionally, several questions were reverse-coded and the concepts of confusion and time pressure were made into a dummy. Next, sentiment

analysis was run with a Python library using TextBlob (Lorai, 2018), converting the answers to the open questions of thoughts to a positive, neutral, or negative loaded number, ranging from one to seven (Gujjar & Kumar, 2021).

After the preparation of the data, SmartPLS 4 (2022) was used to evaluate the measurement model and the structural model (Ringle et al., 2022), since it can deal with complex models with many constructs and model relationships (Hair et al., 2019). A Partial least squares structural equation modeling (PLS-SEM) was run, which combines principal components analysis with ordinary least squares regressions (Hair et al., 2019). Additionally, 5000 bootstrap samples were taken to calculate the significance of the weights, as suggested by Hair et al. (2014).

3.6 Research Ethics and Limitations

To ensure integrity, the researchers adhered to the Netherlands Code of Conduct for Research Integrity, and their principles of honesty, scrupulousness, transparency, independence, and responsibility (KNAW et al., 2018).

Honesty was incorporated by creating an accurate research process, resulting from the use of a checklist for every participant. Also, the data used is pure and the results are not messed with.

Scrupulousness is reflected in a scientifically based research process, based on the approved American Psychology Association (APA) referencing style (Smith, 2003).

Transparency has been taken into account, by being clear about the data gathered by the participant and how this data is used, by letting them sign a consent form. This form includes the right to withdraw from the experience at any time they want and the possibility to be informed about the results of the study.

The independence includes objective and scientific considerations of this research. Data from our respondents will only be used for this research and cannot in any way be linked back to a specific person.

Lastly, this study acknowledges the external influences that could affect the study, therefore taking responsibility. Additionally, responsibility is taken for the storage of personal information. According to Meta, the data from the chatbot, which was run through WhatsApp Business, is stored in Denmark, Ireland, and/or Sweden (*European Data Centres - Meta Data Centers, 2023*).

Nonetheless, this study still has some limitations. Lab-based experiments can be limited in simulating real-life situations. For example, in the consent form, hints are given about eye tracking and expectations of the experiment, which could influence the objectivity of the participant. However, these experiments can be effective for testing causal relationships between variables (Jennings, 2005).

4. Results

A PLS-SEM algorithm was run using the software package SmartPLS 4 (Ringle et al., 2022), to test the hypotheses related to the research question that CX explains the effect of customer confusion on satisfaction and loyalty. A confirmatory factor analysis was run to evaluate the measurement model. Next, the structural model was evaluated.

4.1 Measurement Model

The measurement model deals with the measurement of the latent variables of the PLS-SEM path modeling (Ahmad et al., 2017). By evaluating this model the quality of the constructs is measured. Sensations had been divided into four separate latent variables, based on their measurement level and area of interest. Hence, thoughts and CX outcomes only needed to be

tested for their reliability and validity. Consequently, the factor loadings are assessed, followed by indicating the construct reliability and validity.

The factor loadings indicate how well an item represents the construct it is supposed to measure, expressed between a range from -1.0 to +1.0. A higher loading indicates a higher correlation with the construct (Pett et al., 2003). Outer loadings of 0.50 are acceptable, but values above 0.70 are seen as highly satisfactory (Chin, 1998). Besides the open questions of thoughts, none of the factor loadings were below the value of 0.50 (see Table 1). Nonetheless, two of the items show a value below the threshold, consequently, the lowest value of ‘thoughts open question’ was removed. However, this caused no increase in the value of the second item of the construct, the value declined to 0.445. Additionally, the construct still shows low reliability and validity (CR = 0.660, AVE = 0.521). Moreover, the construct showed low variance explained ($R^2 = 0.057$), hence the decision was made to remove the construct of ‘thoughts open questions’ since it did not meet expectations. The utilitarian value will remain to measure the construct of thoughts.

Table 1

Factor loadings, composite reliability, and average variance extracted from the constructs and their items

Components and items	Loading
Loyalty	<i>CR: 0.811, AVE: 0.688</i>
1. How likely is it that you would tell a friend or colleague about this shopping experience? (Highly Unlikely – Strongly Likely)	0.945*
2. How likely is it that you would use this shopping experience again? (Highly Unlikely – Strongly Likely)	0.694*

Thoughts open questions*CR: 0.608, AVE: 0.363*

- | | |
|---|--------|
| 1. Please take me through what you were thinking as you were attempting to help me resolve the purchase error we encountered. | 0.835* |
| 2. What actions crossed your mind in the process of resolving this error? | 0.447* |
| 3. Would you please elaborate by telling me what was valuable in achieving the outcome of the shopping experience and how? | 0.440* |

Utilitarian value*CR: 0.865, AVE: 0.617*

- | | |
|---|--------|
| 1. To what extent do you agree that we have accomplished what we wanted from this purchase? | 0.722* |
| 2. To what extent do you agree that we were effective in completing the purchase? | 0.771* |
| 3. To what extent do you agree that the way the error was resolved was useful? | 0.751* |
| 4. How strongly do you agree that this shopping experience with me was valuable? | 0.887* |

Note. CR: composite reliability; AVE: average variance extracted; * $p < .01$

Since the construct reliability and validity provide sufficient numbers, the discriminant validity should be looked into as well, as it measures the degree of uniqueness of the concepts (Bagozzi et al., 1991). According to the Fornell and Larcker (1981) criteria, the square root of AVE should be higher than the correlation with all other constructs (see Table F1). This was the case for each construct. Additionally, the Heterotrait-Monotrait Ratio (HTMT) should be lower than .85 (Henseler et al., 2014). The results show a value somewhat above the threshold for loyalty with satisfaction and utilitarian value (see Table F2). Since loyalty had a lack of

discriminant validity, it was decided to remove the lowest scoring item, resulting in a decrease in the HTMT ratio to 0.779 for satisfaction and 0.819 for utilitarian value. Resulting in a higher uniqueness of the concepts and making loyalty valid.

4.1.1 Control Variables

A factor analysis was run for two control variables: tolerance for ambiguity and confusion avoidance. The other control variables were measured with one item. The factor loadings are shown in Table 2. The results show non-significant low loadings, with low reliability and validity. Hence, the second item of tolerance for ambiguity and the first item of confusion avoidance were removed. After removing item 2 of tolerance for ambiguity, item 1 had a loading of 0.877 and item 2 a loading of 0.721 (CR = 0.782, AVE = 0.644). So, the two items measure the construct well. After removing the item of confusion avoidance, the remaining items went up to 0.730 and 0.679 respectively. Nonetheless, the reliability and validity of the construct remained low (CR = 0.664, AVE = 0.497), showing variance between the answers. The discriminant validity shows an overlap of the construct with multiple other ones, such as loyalty (0.945), satisfaction (1.258), and tolerance for ambiguity (4.573). Hence items 1 and 3 were removed, leaving item 2, which showed no overlap with other constructs.

Table 2

Factor loadings, composite reliability, and average variance extracted from the control variables and their items

Components and items	Loading
Tolerance for ambiguity	<i>CR: 0.035, AVE: 0.210</i>
1. To what extent do you agree that you prefer situations in which there is some ambiguity?	-0.264

2. How strongly would you agree that you enjoy tackling problems that are complex enough to be ambiguous?	0.728
3. To what extent do you agree that you generally prefer novelty to familiarity?	-0.171
Confusion avoidance	<i>CR: 0.242, AVE: 0.314</i>
1. To what extent do you agree that providing clear information is the basis for avoiding confusion?	-0.425
2. To what extent do you agree that you want to see unambiguous information when resolving confusing issues occurring in shopping experiences?	0.632
3. To what extent do you agree that you enjoy tackling issues pertaining to confusion occurring from ambiguous information?	0.602

Note. CR: composite reliability; AVE: average variance extracted; *p < .01

4.2 Structural Model

The structural model was evaluated to investigate the links between the different constructs (Hair et al., 2018). To evaluate the hypotheses, the mediation analysis used the bootstrapping procedure with 5,000 resamples to determine the significance of each weight (Hair et al., 2014).

Additionally, the multicollinearity of the model was assessed by looking at the variance inflation factor (VIF). All the predictors of the model were below three, revealing a bias-free model (Knock, 2015). Next, the model fit is assessed. The root mean square residual (SRMR) measures the difference between the observed correlation and the model implied correlation matrix (Ringle et al., 2022). A value below 0.10 is considered a good fit (Hu & Bentler, 1998). The current model (SRMR = 0.104) shows a better fit in comparison to before the deletion of the variables

(SRMR = 0.109), however, still it has a moderate explanatory capacity. Nonetheless, the impact of customer confusion on customer satisfaction and loyalty will be tested, in addition to the mediation research.

The structural model results (see Table F3) show that H1a is statistically significant with a $p < 0.001$, whereas H1b is not. So, customer confusion has a negative significant effect on satisfaction ($R^2 = 0.599$), but not on loyalty ($R^2 = 0.582$). However, both have a large effect size. Moreover, the path coefficients for H2 show only one significant result for the fixations on relevant areas ($p = 0.20$), the other three variables are not supported. So, the effect of customer confusion on gaze behavior is only partially supported. H3 is supported, the negative effect of customer confusion on thoughts is significant ($p < 0.001$). Additionally, thoughts show a medium to large effect size ($R^2 = 0.309$). With concern for the control variables (age, gender, tolerance for ambiguity, confusion avoidance), the results show no significant effect of these variables on satisfaction and loyalty (see Table F4).

An additional analysis was run to test the effect of these control variables on sensations and thoughts, since it is expected to influence the gaze behavior and thoughts of the customer (Furnham & Ribchester, 1995; Simon & Usunier, 2007). Additionally, the effect of alternative explanation was tested, see Table F4. The results show the negative significant effect of age on DNRA ($p = 0.04$), DRA ($p < 0.001$), and thoughts ($p = 0.02$). Secondly, gender (female = 0, male = 1) had a positive significant effect on DNRA ($p < 0.001$) and DRA ($p = 0.009$), and a negative significant effect on FNRA ($p = 0.04$). Lastly, the alternative explanation of the preference for humans when solving problems showed a positive significant effect for FNRA ($p < 0.001$) and a negative effect for thoughts ($p < 0.001$) and satisfaction ($p = 0.04$).

4.2.1 Mediation Analysis

This research expected three different indirect effects (see Table F3). The first indirect effect that was measured was the effect of customer confusion on customer satisfaction and loyalty through sensations. This effect was not significant, since the four variables of sensations showed an insignificant effect ($p > 0.20$). All the effects however were negative. Still, H4a and H4b are not supported. H4a shows no effect at all according to the model of Zhao et al. (2010) since the direct effect on satisfaction was insignificant ($p = 0.43$). H4b shows only a direct effect ($p = 0.03$).

Moreover, the results showed that thoughts significantly mediated the relationship between customer confusion and customer satisfaction ($p < 0.001$) and loyalty ($p < 0.001$), negatively. So, H5a and H5b are supported. Since the direct effect of confusion on satisfaction is insignificant ($p = 0.43$), thoughts fully mediate the relation between confusion and satisfaction. However, the direct effect of confusion on loyalty is significant ($p = 0.03$), making it a competitive partial mediation (Hair et al., 2017; Zhao et al., 2010).

Additionally, a long-way indirect effect of customer confusion on CX outcome, via sensations and thoughts, was expected (H6). This was not supported, since none of the paths through sensations and thoughts showed a significant result on satisfaction and/or loyalty, since all p values were above 0.40. Hence, H6a and H6b are not supported.

4.2.2 Multi-group Analysis with Time Pressure

In the final part of the research, notable variations between the people with and without time pressure were made, in terms of the effect of customer confusion on sensations, thoughts, satisfaction, and loyalty. Only one significant difference was found between people with and without time pressure. The effect of confusion on sensation's DNRA was higher for no pressure ($p = .05$), with a difference of .711. Hence, H7a is partially supported. H7b, c, and d are not.

5. Discussion

The objective of this study was to measure the impact of customer confusion on the CX during the interaction with a chatbot, drawing on the already-established effect of customer confusion on satisfaction and loyalty. To the author's knowledge, this study represents one of the early attempts to investigate the link between customer confusion and CX during a self-service, by suggesting that the CX responses mediate this effect. The findings provide evidence of a mediation effect which proposed that the indirect effect of customer confusion causes more negative thoughts, and lower levels of customer satisfaction and loyalty accordingly.

Additionally, the findings serve as a reminder to companies using chatbots, to listen to their target group. Since, the preference for a human agent can strongly affect the utilitarian value negatively, resulting in dissatisfaction with the service. The key theoretical and managerial implications, limitations, future research, and conclusion are elaborated on next.

5.1 Theoretical Implications

This study adds to the existing knowledge of the effects of customer confusion on the CX. First, although customer confusion is regarded as an antecedent of dissatisfaction and disloyalty (Matzler & Waiguny, 2005), yet according to the literature review, no efforts have been made to study the mediating effect of CX responses in the relationship between customer confusion and CX outcomes. The findings reveal a mediating effect of thoughts in the context of chatbots. In this way, the current study contributes to broadening the concept of CX, which is still uncertain according to Becker and Jaakkola (2020). Suggesting that customer confusion during chatbot interactions negatively impacts the thoughts of a customer. This is in line with the recent CX literature, which revealed that confused individuals have unclear and unstructured thoughts (Leek & Kun, 2006). Also, it provides evidence that the evaluation of service, the utilitarian value,

impacts the satisfaction and loyalty of a customer negatively (Babin et al., 1994; Chiu et al., 2014; Lee & Wu, 2017). Specifically, this study shows that when a customer experiences confusion when shopping with a chatbot, he or she will evaluate the whole shopping experience negatively, resulting in low satisfaction with the purchase and no intention to shop with the bot again.

However, no (mediating) effect was found for the gaze behavior. In contrast, Salminen et al. (2018) claimed that fixations should increase in number and duration, mostly focusing on non-relevant information. The findings of this study show only one positive significant effect of customer confusion on the total fixations in relevant areas. This provides evidence for an increase of fixations, however not in the non-relevant areas. Moreover, no significant relationship was found from confusion on the duration of fixations, it is still interesting to see that people with confusion show shorter durations on the screen and longer on the environment than people without confusion. The latter effect is stronger when time pressure is involved. This is partially in contrast with Pieters and Warlop (1999), who claim that time pressure and confusion will cause a shorter duration of fixations. Hence, the current study provides evidence that the early stream of literature on novel technologies like eye tracking is not coherent, since they provide contrasting results on the effect of confusion on gaze behavior. However, the reliability and validity of the results can be questioned, since the eye tracker was defective and showed low reliability during the experiment.

Second, to understand how the effect of customer confusion on satisfaction and loyalty can be reduced, the moderation of time pressure was studied. The findings showed that time pressure during purchase does not cause stronger effects from confusion. This is in contrast to the study of Xue et al. (2020), who mentions that the lack of time results in overload confusion, followed by a lack of satisfaction with the purchase. However, since the current study provided a

live shopping bot, able to solve the issue on time, this could have reduced the effect of the time pressure (Xue et al., 2020). Additionally, it could be explained since the manipulation of confusion already provoked information overload, hence the addition of time pressure was not significant.

Last, since chatbots and other SSTs have not been around for that long, few studies used this context to study customer confusion (Cunningham et al., 2009; Varela-Neira et al., 2008). Hence, the results of this study provide information on how chatbots can be coupled with the confusion someone experiences. A lack of clear help from the chatbot can lead to confusion, which in turn is blamed on the service provided by the chatbot. Hence, this research shows that dissatisfaction and disloyalty as an effect of customer confusion, can be explained by the thinking process of the customer about the experience and evaluation of the service provided by the chatbot.

5.2 Managerial Implications

Aside from the theoretical contribution, valuable managerial implications can be drawn from the findings of this study. First, as the use of chatbots by firms is rapidly growing, the contribution of this study is valuable (Omar, 2023). When implemented right, the use of chatbots can save firms time and costs, as it replaces employees (Shukairy, 2022). In line with these studies, the current study suggests that firms should implement chatbots with a human-centered design, as a way to cope with customer confusion during a chatbot encounter (Blazevic & Sidaoui, 2020). This study shows that when a customer experiences confusion when shopping with a chatbot, he or she will evaluate the whole shopping experience negatively, resulting in low satisfaction with the purchase and no intention to shop with the bot again. Resulting in the possibility of them shopping somewhere else where their needs are listened to. This should be prevented. Hence it is

important to provide a clear shopping experience and make sure the shopping bot is capable to handle all sorts of problems and learn from feedback. According to the findings, this will potentially increase the satisfaction of the customer with the service since it will provide a more pleasant CX. Moreover, it could enhance satisfaction with the shopping experience provided by the chatbot.

Second, the findings provide evidence for age affecting the relationship between customer confusion and CX. Existing literature revealed that older customers tend to be more negative toward chatbots, whereas the younger generation often prefers it over human interaction (Simon & Usunier, 2007). In line with this literature, the current study showed how older people were generally more negative in their thoughts than the younger generation. Which could be explained by their preference for human interactions. The alternative explanations provided evidence that customers who preferred human agencies for helping to solve problems, were, in general, more negative thinkers and were more dissatisfied. Consequently, marketing managers are advised to take into account their target group when implementing any kind of chatbot. Because when confusion arises during the shopping experience, this can lead to a negative evaluation of the service, which may indirectly damage the share and profitability of the company (Keaveney, 1995). So, even when chatbots are considered more affordable, this does not outweigh the negative effects they can cause on the wrong customer group. Managers need to understand if it is enhancing the CX or not (Walsh & Mitchell, 2010).

Last, successfully integrating chatbots in shopping interactions make it possible to learn from the customer's experience by asking for an evaluation of the service and makes it possible to learn from the customer's feedback, benefitting both the customer and the firm.

5.3 Limitations

Besides the contributions of this study, it also has some limitations. First, the reliability of the results can be questioned as the experience was simulated. No real product or service was bought and the chatbot was operated by a human, so they were not experiencing the real feeling they would have in real life. Also, since it was not clear what they were fictionally buying, this could evoke confusion before it was intentionally evoked. Consequently, this could have affected their evaluation of the shopping experience. Additionally, the shopping environment was restricted, as the participant was not able to freely navigate the internet to search for additional information or refresh the page. Lowering the feeling of a real event. Moreover, the effect of the error page and the timer might not be so effective as in real life, since people did not have the fear of losing their purchase which they would have if they had the motivation to buy a certain product or service. All these different factors may have decreased the relationships in the model, then when having a real shopping experience.

Second, as CX is a complex and multidimensional phenomenon, this study only looked at a part of the CX (Lemon & Verhoef, 2016; Schmitt, 1999). Therefore, no generalizations can be made about the effect of customer confusion on the other CX responses. Also, as no significant results were found on the sensations of the customer, it cannot be expected that all the responses of a customer are influenced by someone's state of confusion.

Third, the lack of significant results and data on the sensations concept can be explained by the partial defect eye tracker used during the experiment. Since one of the cameras was not working, this could have affected the results, and the validity and reliability of the research.

Fourth, although the sample size of the study was sufficient, with an eventual number of 197 participants, it was not a diverse group of people. Only the students and employees of the

Radboud University were asked to participate, with an occasional student from a different location, ranging from 18 to 61 years old. As a consequence, the results may be biased since the population had an average age of 25 years old and were highly educated, this is not representable for the whole population.

Last, the number of items used to measure satisfaction, loyalty, and confusion avoidance, were limited to one. Since these measures were reduced to one item and adapted to the environment of the experiment, the questions were not tested before in existing literature. This could explain the insignificant result of customer confusion on loyalty, as this relationship was yet established in the literature (Matzler & Waiguny, 2005). Moreover, since this study is not longitudinal, the long-term behavior of the participant could not be measured efficiently, such as loyalty. This could explain the insufficient results with this concept (Griffin, 1995). In addition, the open-question items of thoughts had their limitations as well, hence the deletion of them. The sentiment analysis used to analyze the open questions of thoughts could only evaluate the answers by giving them a negative, neutral, or positive label. Since thoughts also express underlying reasons and intentions and are more than just positive or negative feelings, this could not be taken into account. Moreover, the quantitative nature of this study made it difficult to include these open answers in the analysis, as it went more exploratory than the other constructs, making it hard to analyze their opinions individually. These limitations concerning the measurement of the concepts could explain the low model fit.

5.4 Future Research

Drawing on the discussion and limitations above, there are a few topics that need further research. First, to make the result of this study more valid and reliable, future research should replicate this study in a real setting. Instead of a wizard-of-oz setting where a human was

communicating with the participant, a fully automated chatbot should be used. Additionally, applying this research in a real setting, where an actual product or service should be bought, would make the result more valid. This will limit the doubts and confusion regarding the shopping experience itself, which could influence the manipulated confusion condition.

Second, conducting this study with different demographics would greatly benefit the current study. The findings show a difference for people who are older regarding their thoughts and CX outcomes. However, only a small portion of the participant was above 50 and only one above 60, hence it could be interesting if this result remains for an older generation. Additionally, different education levels could affect the current results. This could influence their thinking process and problem-solving skills. Testing this study on a varied population will make the results more generalizable.

Third, it is important to assess whether the results for sensations stay the same when measured with more appropriate eye tracking technology. Since gaze behavior was not sufficiently measured in this study, future research could consider using other eye tracking techniques and software to measure the sensations of a person in a state of confusion. Since this study did not find any significant effects, this should be tested through different studies. Also, other sensations could be taken into account, like sense, touch, and sound (Schmitt, 1999). Moreover, different CX responses in addition to sensations and thoughts should be considered to provide a clear picture of how CX mediates the relationship between customer confusion and CX outcomes.

Lastly, to be able to understand the current results better, a qualitative study including the indirect relationship through thoughts could be considered. Since the current study was not able to include the thought process of the customer, analyzing the data qualitatively would make it

possible to understand the reasons why the customer's thoughts were affected and why this made them dissatisfied and disloyal.

5.5 Conclusion

This study looked at the impact of customer confusion on the CX with a shopping bot. Based on a wizard-of-oz experiment and eye tracking data, results suggest that the presence of thoughts as a mediator reveals that customer confusion negatively impacts utilitarian value. Subsequently, this negative effect on thoughts leads to dissatisfaction and disloyalty. However, no significant effect was found on the eye sensations of customers. Moving forward, firms can use these insights to improve their chatbots and raise the overall CX. Nonetheless, the study was limited by experimental design, which should be taken into consideration when implementing the results. Consequently, future research could replicate this study in a real shopping setting, exploring the CX more in-depth.

References

- Ahmed, S., Ahmad, F. B., & Jaaffar, A. R. (2017). Influence of employee engagement on employee promotion opportunity and performance relationship in developing context: Critical evaluation with PLS-SEM analysis technique. *Transylvanian Review*, 25(17), 4327-4340.
- Almeida, F., Santos, J. D., & Monteiro, J. A. (2020). The challenges and opportunities in the digitalization of companies in a post-COVID-19 World. *IEEE Engineering Management Review*, 48(3), 97-103. <https://doi.org/10.1109/EMR.2020.3013206>
- Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: Findings from Sweden. *Journal of Marketing*, 58(3), 53-66. <https://doi.org/10.1177/002224299405800304>
- Anninou, I., & Foxall, G. R. (2019). The reinforcing and aversive consequences of customer experience. The role of consumer confusion. *Journal of retailing and consumer services*, 51, 139-151. <https://doi.org/10.1016/j.jretconser.2019.05.029>
- Ares, G., Giménez, A., Bruzzone, F., Vidal, L., Antúnez, L., & Maiche, A. (2013). Consumer Visual Processing of Food Labels: Results from an Eye-Tracking Study. *Journal of Sensory Studies*, 28(2), 138–153. <https://doi.org/10.1111/joss.12031>
- Ashby, N. J. S., Johnson, J., Krajbich, I., & Wedel, M. (2016). Applications and Innovations of Eye-movement Research in Judgment and Decision Making. *Journal of Behavioral Decision Making*, 29(2–3), 96–102. <https://doi.org/10.1002/bdm.1956>
- Babin, B. J., Darden, W. R., & Griffin, M. (1994). Work and/or Fun: Measuring Hedonic and Utilitarian Shopping Value. *Journal of Consumer Research*, 20(4), 644. <https://doi.org/10.1086/209376>

- Bagherzadeh, R., Rawal, M., Wei, S., & Torres, J. (2020). The journey from customer participation in service failure to co-creation in service recovery. *Journal of Retailing and Consumer Services*, 54, 102058. <https://doi.org/10.1016/j.jretconser.2020.102058>
- Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing Construct Validity in Organizational Research. *Administrative Science Quarterly*, 36(3), 421. <https://doi.org/10.2307/2393203>
- Beatson, A., Lee, N., & Coote, L. V. (2007). Self-service technology and the service encounter. *The Service Industries Journal*, 27(1), 75-89. <https://doi.org/10.1080/02642060601038700>
- Becker, L., & Jaakkola, E. (2020). Customer experience: fundamental premises and implications for research. *Journal of the Academy of Marketing Science*, 48, 630-648. <https://doi.org/10.1007/s11747-019-00718-x>
- Bell, C. R., & Zemke, R. E. (1987). Service breakdown: the road to recovery. *Management review*, 76(10), 32.
- Blazevic, V., & Sidaoui, K. (2022). The TRISEC framework for optimizing conversational agent design across search, experience and credence service contexts. *Journal of Service Management*, 33(4/5), 733–746. <https://doi.org/10.1108/josm-10-2021-0402>
- Bustamante, J. M., & Rubio, N. (2017). Measuring customer experience in physical retail environments. *Journal of Service Management*, 28(5), 884–913. <https://doi.org/10.1108/josm-06-2016-0142>
- Bowen, J. R., & Chen, S. (2001). The relationship between customer loyalty and customer satisfaction. *International Journal of Contemporary Hospitality Management*, 13(5), 213–217. <https://doi.org/10.1108/09596110110395893>

- Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern methods for business research*, 295(2), 295-336. <https://psycnet.apa.org/record/1998-07269-010>
- Chiu, C., Wang, E. W., Fang, Y., & Huang, H. (2012). Understanding customers' repeat purchase intentions in B2C e-commerce: the roles of utilitarian value, hedonic value and perceived risk. *Information Systems Journal*, 24(1), 85–114. <https://doi.org/10.1111/j.1365-2575.2012.00407.x>
- Collier, J. H., Moore, R. B., Horkey, A., & Moore, M. J. (2015). Why the little things matter: Exploring situational influences on customers' self-service technology decisions. *Journal of Business Research*, 68(3), 703–710. <https://doi.org/10.1016/j.jbusres.2014.08.001>
- Connor, C. B., Egeth, H. E., & Yantis, S. (2004). Visual Attention: Bottom-Up Versus Top-Down. *Current Biology*, 14(19), R850–R852. <https://doi.org/10.1016/j.cub.2004.09.041>
- Cunningham, L. A., Young, C. E., & Gerlach, J. H. (2009). A comparison of consumer views of traditional services and self-service technologies. *Journal of Services Marketing*, 23(1), 11–23. <https://doi.org/10.1108/08876040910933057>
- Curran, J., Meuter, M. L., & Surprenant, C. F. (2003). Intentions to Use Self-Service Technologies: A Confluence of Multiple Attitudes. *Journal of Service Research*, 5(3), 209–224. <https://doi.org/10.1177/1094670502238916>
- Dabholkar, P. A., & Spaid, B. I. (2012). Service failure and recovery in using technology-based self-service: effects on user attributions and satisfaction. *The Service Industries Journal*, 32(9), 1415-1432. <https://doi.org/10.1080/02642069.2011.600518>
- Dean, D. H. (2008). Shopper age and the use of self-service technologies. *Managing Service Quality*, 18(3), 225–238. <https://doi.org/10.1108/09604520810871856>

- Durach, D., & Diachenko, A. (2023, March 1). *The future of chatbots: 10 trends, latest stats & market size*. Onix Systems. Retrieved March 13, 2023, from <https://onix-systems.com/blog/6-chatbot-trends-that-are-bringing-the-future-closer>
- Edward, M., & Sahadev, S. (2012). Modeling the consequences of customer confusion in a service marketing context: An empirical study. *Journal of Services Research*, 12(2).
- Ellis, R. S., & Allaire, J. C. (1999). Modeling Computer Interest in Older Adults: The Role of Age, Education, Computer Knowledge, and Computer Anxiety. *Human Factors*, 41(3), 345–355. <https://doi.org/10.1518/001872099779610996>
- European Data Centres - Meta Data Centers*. (2023, May 29). Meta Data Centers. <https://datacenters.atmeta.com/europe/>
- Feine, J., Morana, S., & Gnewuch, U. (2019). Measuring Service Encounter Satisfaction with Customer Service Chatbots using Sentiment Analysis. *Wirtschaftsinformatik Und Angewandte Informatik*, 1115–1129. <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1248&context=wi2019>
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding and Evaluating Qualitative Research. *Australian & New Zealand Journal of Psychiatry*, 36(6), 717–732. <https://doi.org/10.1046/j.1440-1614.2002.01100.x>
- Foxman, E. R., Berger, P. W., & Cote, J. A. (1992). Consumer brand confusion: A conceptual framework. *Psychology & Marketing*, 9(2), 123-141. <https://doi.org/10.1002/mar.4220090204>

- Foxman, E. F., Muehling, D. D., & Berger, P. (1990). An Investigation of Factors Contributing to Consumer Brand Confusion. *Journal of Consumer Affairs*, 24(1), 170–189.
<https://doi.org/10.1111/j.1745-6606.1990.tb00264.x>
- Furnham, A., & Ribchester, T. (1995). Tolerance of ambiguity: A review of the concept, its measurement and applications. *Current Psychological Research & Reviews*, 14(3), 179–199. <https://doi.org/10.1007/bf02686907>
- Ganti, A. (2023, March 11). *Central limit theorem (CLT): Definition and key characteristics*. Investopedia. Retrieved April 23, 2023, from https://www.investopedia.com/terms/c/central_limit_theorem.asp
- Gartner. (2022, July 27). *Gartner predicts Chatbots will become a primary customer service channel within five years*. Retrieved March 12, 2023, from <https://www.gartner.com/en/newsroom/press-releases/2022-07-27-gartner-predicts-chatbots-will-become-a-primary-customer-service-channel-within-five-years>
- Goeschl, T., & Lohse, J. (2018). Cooperation in public good games. Calculated or confused? *European Economic Review*, 107, 185–203.
<https://doi.org/10.1016/j.eurocorev.2018.05.007>
- Grégoire, Y., & Mattila, A. S. (2020). Service Failure and Recovery at the Crossroads: Recommendations to Revitalize the Field and its Influence. *Journal of Service Research*, 109467052095807. <https://doi.org/10.1177/1094670520958073>
- Griffin, J. G. (1995). Customer loyalty: how to earn it, how to keep it. *Choice Reviews Online*, 32(11), 32–6312. <https://doi.org/10.5860/choice.32-6312>
- Gujjar, J. P., & Kumar, H. P. (2021). Sentiment analysis: Textblob for decision making. *Int. J. Sci. Res. Eng. Trends*, 7(2), 1097-1099.

- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2018). The results of PLS-SEM article information. *European Business Review*, *31*(1), 2-24.
- Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. <https://doi.org/10.1108/eb-11-2018-0203>
- Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM). *European Business Review*, *26*(2), 106–121. <https://doi.org/10.1108/eb-10-2013-0128>
- Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2017). *Advanced Issues in Partial Least Squares Structural Equation Modeling*. https://openlibrary.org/books/OL26377484M/Advanced_issues_in_partial_least_squares_structural_equation_modeling
- Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Hess, R. A., Ganesan, S., & Klein, N. M. (2003). Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction. *Journal of the Academy of Marketing Science*, *31*(2), 127–145. <https://doi.org/10.1177/0092070302250898>
- Hoque, M. Z., Bashar, M. K., & Akhter, F. (2021). Ambiguity tolerance and confusion avoidance in the intent to purchase farmed fish. *Journal of Applied Aquaculture*, *34*(3), 564–597. <https://doi.org/10.1080/10454438.2021.1872462>
- Hoyer, W. D., Kroschke, M., Schmitt, B. H., Kraume, K., & Shankar, V. (2020). Transforming the Customer Experience Through New Technologies. *Journal of Interactive Marketing*, *51*, 57–71. <https://doi.org/10.1016/j.intmar.2020.04.001>

- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037/1082-989x.3.4.424>
- IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp
- Innis, D. E., & La Londe, B. J. (1994). Customer service: the key to customer satisfaction, customer loyalty, and market share. *Journal of business Logistics*, 15(1), 1. <https://doi.org/1994-01-010735-3766>
- Jennings, G. R. (2005). Business, social science methods used in. *Encyclopedia of Social Measurement*, 1, 219–230. <https://doi.org/10.1016/b0-12-369398-5/00270-x>
- Keaveney, S. M. (1995). Customer Switching Behavior in Service Industries: An Exploratory Study. *Journal of Marketing*, 59(2), 71. <https://doi.org/10.2307/1252074>
- Kaczorowska-Spychalska, D. (2019). How chatbots influence marketing. *Management*, 23(1), 251–270. <https://doi.org/10.2478/manment-2019-0015>
- Keiningham, T. L., Ball, J., Benoit, S., Bruce, H., Buoye, A., Dzenkovska, J., Nasr, L., Ou, Y., & Zaki, M. J. (2017). The interplay of customer experience and commitment. *Journal of Services Marketing*, 31(2), 148–160. <https://doi.org/10.1108/jsm-09-2016-0337>
- KNAW, NFU, NWO, TO2-federatie, Vereniging Hogescholen, & VSNU. (2018). Netherlands Code of Conduct for Research Integrity. <https://doi.org/10.17026/dans-2cijnvw>
- Kock, N. (2015). Common Method Bias in PLS-SEM. *International Journal of e-Collaboration*, 11(4), 1–10. <https://doi.org/10.4018/ijec.2015100101>

- Kwak, S. K., & Kim, J. M. (2017). Central limit theorem: the cornerstone of modern statistics. *Korean Journal of Anesthesiology*, 70(2), 144. <https://doi.org/10.4097/kjae.2017.70.2.144>
- Lee, C. H., & Wu, J. J. (2017). Consumer online flow experience: The relationship between utilitarian and hedonic value, satisfaction and unplanned purchase. *Industrial Management & Data Systems*, 117(10), 2452-2467. <https://doi.org/10.1108/IMDS-11-2016-0500>
- Leek, S., & Kun, D. (2006). Consumer confusion in the Chinese personal computer market. *Journal of Product & Brand Management*, 15(3), 184–193. <https://doi.org/10.1108/10610420610668621>
- Leiva, F. M., Rios, F. J. M., & Martínez, T. L. (2006). Assessment of interjudge reliability in the open-ended questions coding process. *Quality and Quantity*, 40, 519-537. <https://doi.org/10.1007/s11135-005-1093-6>
- Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of marketing*, 80(6), 69-96. <https://doi.org/10.1509/jm.15.0420>
- Lewis, B. R., & Mitchell, V. W. (1990). Defining and measuring the quality of customer service. *Marketing intelligence & planning*, 8(6), 11-17. <https://doi.org/10.1108/EUM0000000001086>
- Loria, S. (2018). textblob Documentation. *Release 0.15*, 2.
- Luarn, P., & Lin, H. (2003). A Customer Loyalty Model for E-Service Context. *Journal of Electronic Commerce Research*, 4, 156–167. <https://www.csulb.edu/web/journals/jecr/issues/20034/Paper3.pdf>

Malhotra, N., Hall, J., Shaw, M., & Oppenheim, P. (2006). *Marketing research : an applied orientation* (Version 1). Deakin University.

<https://hdl.handle.net/10536/DRO/DU:30010407> ([])

Mannheim, I., Cohn-Schwartz, E., Xi, W., Buttigieg, S. C., McDonnell-Naughton, M., Wouters, E., & Van Zaalen, Y. (2019). Inclusion of Older Adults in the Research and Design of Digital Technology. *International Journal of Environmental Research and Public Health*, *16*(19), 3718. <https://doi.org/10.3390/ijerph16193718>

Maruping, L. M., Venkatesh, V., Thatcher, S. M. B., & Patel, P. C. (2015). Folding Under Pressure or Rising to the Occasion? Perceived Time Pressure and the Moderating Role of Team Temporal Leadership. *Academy of Management Journal*, *58*(5), 1313–1333. <https://doi.org/10.5465/amj.2012.0468>

Matzler, K., & Waiguny, M. K. J. (2005). Consequences of Customer Confusion in Online Hotel Booking. In *Springer eBooks*. https://doi.org/10.1007/3-211-27283-6_28

McCullough, M. A., Berry, L. L., & Yadav, M. S. (2000). An Empirical Investigation of Customer Satisfaction after Service Failure and Recovery. *Journal of Service Research*, *3*(2), 121–137. <https://doi.org/10.1177/109467050032002>

McLain, D. L. (2009). Evidence of the Properties of an Ambiguity Tolerance Measure: The Multiple Stimulus Types Ambiguity Tolerance Scale–II (MSTAT–II). *Psychological Reports*, *105*(3), 975–988. <https://doi.org/10.2466/pr0.105.3.975-988>

McLain, D. L., Kefallonitis, E., & Armani, K. (2015). Ambiguity tolerance in organizations: definitional clarification and perspectives on future research. *Frontiers in Psychology*, *6*. <https://doi.org/10.3389/fpsyg.2015.00344>

- Meißner, M., & Oll, J. (2019). The promise of eye-tracking methodology in organizational research: A taxonomy, review, and future avenues. *Organizational Research Methods*, 22(2), 590-617. <https://doi.org/10.1177/1094428117744882>
- Meuter, M. L., Bitner, M. J., Ostrom, A. L., & Brown, S. W. (2005). Choosing among alternative service delivery modes: An investigation of customer trial of self-service technologies. *Journal of Marketing*, 69(2), 61-83. <https://doi.org/10.1509/jmkg.69.2.61.60759>
- Mitchell, V. W., & Papavassiliou, V. (1999). Marketing causes and implications of consumer confusion. *Journal of Product & Brand Management*, 8(4), 319-342. <https://doi.org/10.1108/10610429910284300>
- Mitchell, V., Walsh, G., & Yamin, M. (2005). Towards a Conceptual Model of Consumer Confusion. *Advances in Consumer Research*, 32, 143–150. http://www.acrwebsite.org/volumes/v32/acr_vol32_68.pdf
- Mohammed, S., & Nadkarni, S. (2011). Temporal Diversity and Team Performance: The Moderating Role of Team Temporal Leadership. *Academy of Management Journal*, 54(3), 489–508. <https://doi.org/10.5465/amj.2011.61967991>
- Moore, T., & Zirnsak, M. (2017). Neural Mechanisms of Selective Visual Attention. *Annual Review of Psychology*, 68(1), 47–72. <https://doi.org/10.1146/annurev-psych-122414-033400>
- Morris, M. A., & Venkatesh, V. (2000). Age differences in technology adoption decisions: Implications for a changing work force. *Personnel Psychology*, 53(2), 375–403. <https://doi.org/10.1111/j.1744-6570.2000.tb00206.x>

- Ok, C., Back, K., & Shanklin, C. W. (2005). Modeling Roles of Service Recovery Strategy: A Relationship-Focused View. *Journal of Hospitality & Tourism Research*, 29(4), 484–507. <https://doi.org/10.1177/1096348005276935>
- Oliver, R. P. (1980). A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions. *Journal of Marketing Research*, 17(4), 460. <https://doi.org/10.2307/3150499>
- Omar. (2023, June 17). 23+ Chatbot Statistics 2023: Usage, Benefits & Features. *OJ Digital Solutions*. <https://ojdigitalsolutions.com/chatbot-statistics/>
- Pachman, M., Arguel, A., Lockyer, L., Kennedy, G., & Lodge, J. (2016). Eye tracking and early detection of confusion in digital learning environments: Proof of concept. *Australasian Journal of Educational Technology*, 32(6). <https://doi.org/10.14742/ajet.3060>
- Parasuraman, A. (1996, October). Understanding and leveraging the role of customer service in external, interactive and internal marketing. In *Frontiers in Services Conference, Nashville, TN* (No. 1996/10, p. 3).
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). Making sense of factor analysis: *The use of factor analysis for instrument development in health care research..* <http://ci.nii.ac.jp/ncid/BA62802887>
- Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing*, 16(1), 1–16. [https://doi.org/10.1016/s0167-8116\(98\)00022-6](https://doi.org/10.1016/s0167-8116(98)00022-6)
- Pupil Labs. (2023). Pupil Labs v3.5.1 [Computer software]. Pupil Labs. Retrieved from <https://www.pupil-labs.com/>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, 81(12), 46-55.

Riek, L. D. (2012). Wizard of oz studies in hri: a systematic review and new reporting guidelines. *Journal of Human-Robot Interaction*, 1(1), 119-136.
<https://doi.org/10.5898/jhri.1.1.riek>

Ringle, C. M., Wende, S., & Becke, J.-M. (2022). *Model fit*. SmartPLS 4.
<https://www.smartpls.com/documentation/algorithms-and-techniques/model-fit/>

Ross, A., & Willson, V. L. (2017). Independent Samples T-Test. In *SensePublishers eBooks* (pp. 13–16). https://doi.org/10.1007/978-94-6351-086-8_3

Roth, P. L., & Switzer, F. S. (1995). A Monte Carlo Analysis of Missing Data Techniques in a HRM Setting. *Journal of Management*, 21(5), 1003–1023.
<https://doi.org/10.1177/014920639502100511>

Rowley, J. (2000). Product searching with shopping bots. *Internet Research*, 10(3), 203-214.
<https://doi.org/10.1108/10662240010331957>

Rozin, P., & Cohen, A. B. (2003). High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans. *Emotion*, 3(1), 68-75. <https://doi.org/10.1037/1528-3542.3.1.68>

Salminen, J., Jansen, B. J., An, J., Jung, S.-G., Nielsen, L., & Kwak, H. (2018). Fixation and confusion: investigating eye-tracking participant's exposure to information in personas.

Proceedings of the 2018 Conference on Human Information Interaction & Retrieval - CHIIR '18. <https://doi.org/10.1145/3176349.3176391>.

Salvucci, D. D. (1999). *Mapping eye movements to cognitive processes*. Carnegie Mellon University.

Smith, D. (2003, January). *Five principles for research ethics*. American Psychology Association. Retrieved April 23, 2023, from <https://www-apa-org.ru.idm.oclc.org/monitor/jan03/principles>

Schmitt, B. H. (1999). Experimental marketing: how to get customers to sense, feel, think, act and relate to your company and brands. *European Management Journal*, 18(6), 695.

Schweizer, M., & Kotouc, A. J. (2006). Scale Development for Consumer Confusion. *ACR North American Advances*. http://www.acrwebsite.org/volumes/v33/v33_55.pdf

Shukairy, A. (2022, May 4). *Chatbots In Customer Service – Statistics and Trends [Infographic]*. Invesp. <https://www.invespro.com/blog/chatbots-customer-service/>

Sidaoui, K., Jaakkola, M., & Burton, J. (2020). AI feel you: customer experience assessment via chatbot interviews. *Journal of Service Management*, 31(4), 745-766.
<https://doi.org/10.1108/JOSM-11-2019-0341>

Simon, F., & Usunier, J. (2007). Cognitive, demographic, and situational determinants of service customer preference for personnel-in-contact over self-service technology. *International Journal of Research in Marketing*, 24(2), 163–173.
<https://doi.org/10.1016/j.ijresmar.2006.11.004>

Singh, H. (2006). The importance of customer satisfaction in relation to customer loyalty and retention. *Academy of Marketing Science*, 60(193-225), 46. <https://www.van-haaften.nl/images/documents/pdf/The%20Importance%20of%20Customer%20Satisfaction%20in%20Relation%20to%20Customer%20Loyalty%20and%20Retention.pdf>

- Smith, A. K., & Bolton, R. N. (1998). An experimental investigation of customer reactions to service failure and recovery encounters: paradox or peril?. *Journal of service research, 1*(1), 65-81. <https://doi.org/10.1177/109467059800100106>
- Spreng, R. A., Harrell, G. D., & Mackoy, R. (1995). Service recovery: Impact on satisfaction and intentions. *Journal of Services Marketing, 9*(1), 15–23. <https://doi.org/10.1108/08876049510079853>
- Sproles, G. B., & Kendall, E. (1986). A Methodology for Profiling Consumers' Decision-Making Styles. *Journal of Consumer Affairs, 20*(2), 267–279. <https://doi.org/10.1111/j.1745-6606.1986.tb00382.x>
- Tessier-Lavigne, M. (2000). Visual processing by the retina. *Principles of neural science, 507-522*.
- Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica, 135*(2), 77–99. <https://doi.org/10.1016/j.actpsy.2010.02.006>
- Theron, E., & Terblanche, N. S. (2010). Dimensions of Relationship Marketing in Business-to-Business Financial Services. *International Journal of Market Research, 52*(3), 373–392. <https://doi.org/10.2501/s1470785310201326>
- Thomas, L., & Lleras, A. (2007). Moving eyes and moving thought: On the spatial compatibility between eye movements and cognition. *Psychonomic Bulletin & Review, 14*(4), 663–668. <https://doi.org/10.3758/bf03196818>
- Tracy, B. (1998). E-tailing: what web customers really want. *Advertising Age's Business Marketing, 83*(11), 39-41.
- Turnbull, P. W., Leek, S., & Ying, G. (2000). Customer confusion: The mobile phone market. *Journal of Marketing Management, 16*(1-3), 143-163. <https://doi.org/10.1362/026725700785100523>

- Varela-Neira, C., Vázquez-Casielles, R., & Iglesias-Argüelles, V. (2008). The influence of emotions on customer's cognitive evaluations and satisfaction in a service failure and recovery context. *Service Industries Journal*, 28(4), 497–512.
<https://doi.org/10.1080/02642060801917612>
- Venetis, K., & Ghauri, P. N. (2004). Service quality and customer retention: building long-term relationships. *European Journal of Marketing*, 38(11/12), 1577–1598.
<https://doi.org/10.1108/03090560410560254>
- Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *Management Information Systems Quarterly*, 36(1), 157.
<https://doi.org/10.2307/41410412>
- Verma, S., & Gustafsson, A. (2020). Investigating the emerging COVID-19 research trends in the field of business and management: A bibliometric analysis approach. *Journal of Business Research*, 118, 253-261. <https://doi.org/10.1016/j.jbusres.2020.06.057>
- Verma, S. K., Sharma, R., Deb, S., & Maitra, D. (2021). Artificial intelligence in marketing: Systematic review and future research direction. *International Journal of Information Management Data Insights*, 1(1), 100002. <https://doi.org/10.1016/j.ijime.2020.100002>
- Wachowiak, L., Tisnikar, P., Canal, G., Coles, A., Leonetti, M., & Celiktutan, O. (2022). Analysing Eye Gaze Patterns during Confusion and Errors in Human–Agent Collaborations. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. <https://doi.org/10.1109/ro-man53752.2022.9900589>

- Walsh, G., & Mitchell, V. W. (2010). The effect of consumer confusion proneness on word of mouth, trust, and customer satisfaction. *European Journal of marketing*, 44(6), 838-859. <https://doi.org/10.1108/03090561011032739>
- Wedel, M., & Pieters, R. (2008). A Review of Eye-Tracking Research in Marketing. In *Emerald Group Publishing Limited eBooks* (pp. 123–147). [https://doi.org/10.1108/s1548-6435\(2008\)0000004009](https://doi.org/10.1108/s1548-6435(2008)0000004009)
- Wobker, I., Eberhardt, T., & Kenning, P. (2015). Consumer confusion in German food retailing: the moderating role of trust. *International Journal of Retail & Distribution Management*, 43(8), 752–774. <https://doi.org/10.1108/ijrdm-07-2012-0061>
- Xue, P., Jo, W., & Bonn, M. A. (2020). Online hotel booking decisions based on price complexity, alternative attractiveness, and confusion. *Journal of Hospitality and Tourism Management*, 45, 162–171. <https://doi.org/10.1016/j.jhtm.2020.08.013>
- Yang, J., & Klassen, K. J. (2008). How financial markets reflect the benefits of self-service technologies. *Journal of Enterprise Information Management*, 21(5), 448–467. <https://doi.org/10.1108/17410390810904238>
- Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>
- Ziv, G. (2016). Gaze Behavior and Visual Attention: A Review of Eye Tracking Studies in Aviation. *The International Journal of Aviation Psychology*, 26(3–4), 75–104. <https://doi.org/10.1080/10508414.2017.1313096>

Appendix A. Manipulation

Table A1. *Independent sample t-test outcome*

Construct	Items	No		Yes		p
		M	SD	M	SD	
Confusion	How strongly do you agree that the more information you sought, the harder it seemed to complete the purchase?	3.60	1.75	5.24	1.30	<.001
	To what extent do you agree that the provided information was clear enough to complete the purchase?	2.85	1.89	4.94	1.73	<.001
	How strongly do you agree with: the information provided was so ambiguous that you often felt confused?	3.53	1.96	5.74	1.32	<.001
	To what extent do you agree: that seeking more information made completing your purchase less confusing?	3.13	1.55	4.13	1.50	<.001
Time Pressure	To what extent do you agree that you had enough time to complete the purchase?	2.52	1.50	2.67	1.54	.234
	How strongly do you agree that you felt like you were under time pressure while making the purchase?	3.51	1.88	4.21	2.02	<.05

To what extent do you agree 3.74 1.86 5.11 1.84 <.001
that the presence of a timer
made you feel rushed during
the purchase?

Note: M = Mean, SD = Standard Deviation.

Figure A1. *Condition one (no confusion and no time pressure)*

Elite Forté

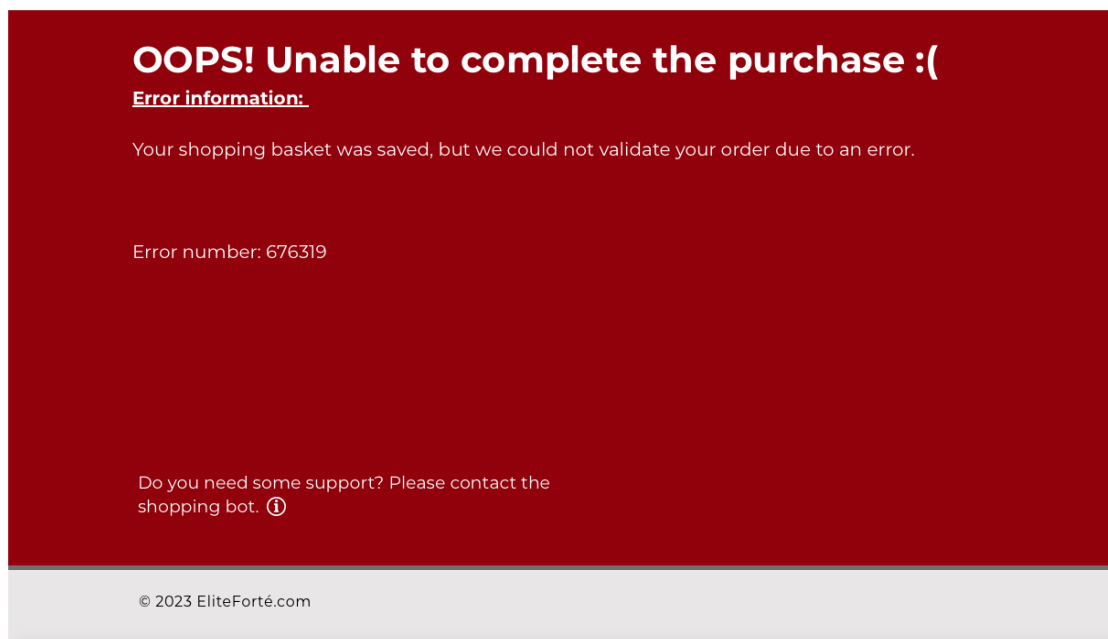


Figure A2. *Condition two (yes confusion and no time pressure)*

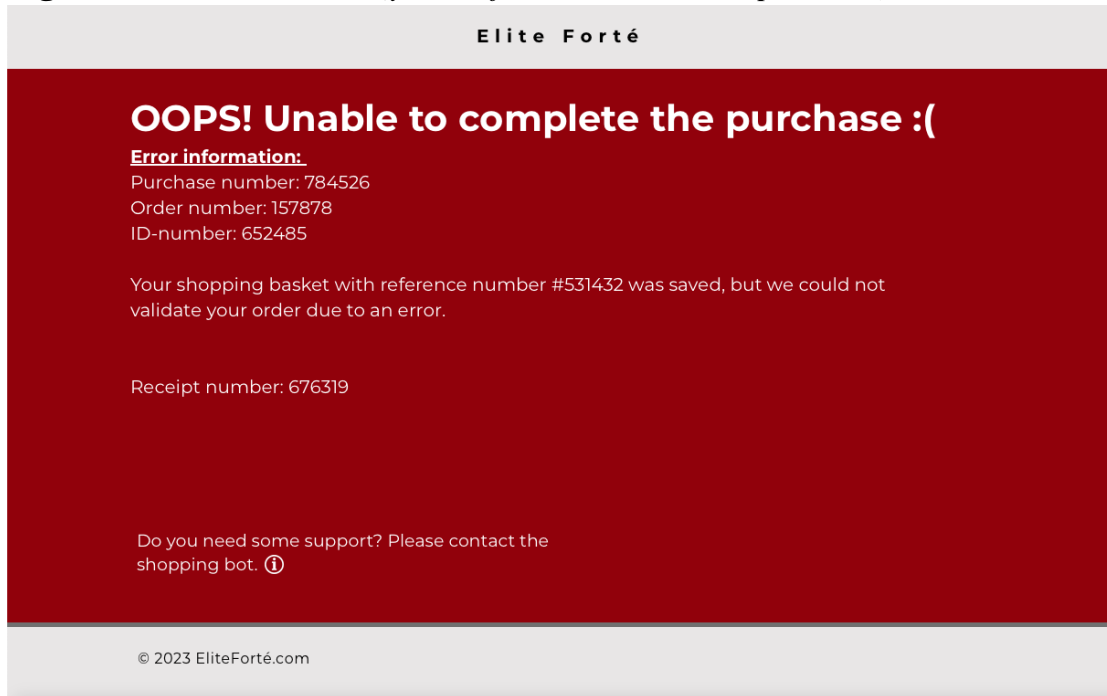


Figure A3. *Condition three (no confusion and yes time pressure)*

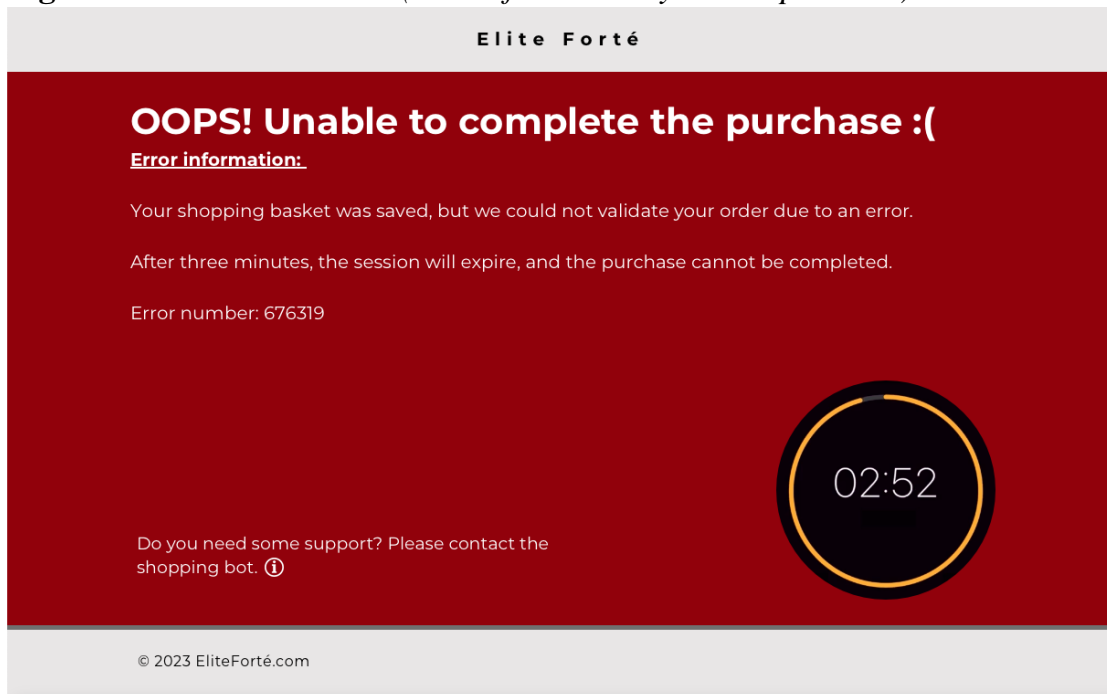


Figure A4. *Condition four (yes confusion and yes time pressure)*


Elite Forté

OOPS! Unable to complete the purchase :(


Error information:
Purchase number: 784526
Order number: 157878
ID-number: 652485

Your shopping basket with reference number #531432 was saved, but we could not validate your order due to an error. After three minutes, the session will expire, and the purchase cannot be completed

Receipt number: 676319



02:53

Do you need some support? Please contact the shopping bot. 

© 2023 EliteForté.com

Appendix B. Operationalization Table

Concept	Measurement description	Source
Sensations		
<i>Gaze behavior</i>	Total number of fixations	Salminen et al. (2018)
	Average duration of fixations	
Thoughts		
	<p>Please take me through what you were thinking as you were attempting to help me resolve the purchase error we encountered.</p> <p>What actions crossed your mind in the process of resolving this error?</p> <p>Would you please elaborate by telling me what was valuable in achieving the outcome of the shopping experience and how?</p>	-
<i>Utilitarian value</i>	To what extent do you agree that we have accomplished what we wanted from this purchase?	Babin et al. (1994)
	To what extent do you agree that we were effective in completing the purchase?	
	To what extent do you agree that the way the error was resolved was useful?	
	How strongly do you agree that this shopping experience with me was valuable?	

**Customer
experience**

outcomes

<i>Satisfaction</i>	So as an outcome, to what extent would you say that you are satisfied with the shopping experience? (Strongly Dissatisfied – Strongly Satisfied)	-
<i>Loyalty</i>	How likely is it that you would tell a friend or colleague about this shopping experience? (Highly Unlikely – Strongly Likely) How likely is it that you would use this shopping experience again? (Highly Unlikely – Strongly Likely)	Reichheld (2003)

**Control
variables**

<i>Age</i>	Can you please provide me with your age?	-
<i>Gender</i>	What is the gender you identify with the most? (1) Male, (2) Female, (3) Other, (4) Prefer not to say	-
<i>Tolerance for ambiguity</i>	To what extent do you agree that you prefer situations in which there is some ambiguity?	McLain (2009)

How strongly would you agree that you enjoy tackling problems that are complex enough to be ambiguous?

To what extent do you agree that you generally prefer novelty to familiarity?

Confusion To what extent do you agree that providing clear information is the basis for avoiding confusion? Schweizer et al. (2006)

Avoidance To what extent do you agree that you want to see unambiguous information when resolving confusing issues occurring in shopping experiences?

To what extent do you agree that you enjoy tackling issues pertaining to confusion occurring from ambiguous information?

Realness of chatbot To what extent do you agree that I behaved like an automated assistant? -

Distraction To what extent do you agree that you sensed something in your surroundings that hindered you from fully concentrating on your experience with me? -

Problem solving preference To what extent do you agree that you would have preferred chatting to a human agent as -

opposed to me when encountering the error we had?

Manipulation**Check**

<i>Confusion</i>	How strongly do you agree that the more information you sought, the harder it seemed to complete the purchase? To what extent do you agree that the provided information was clear enough to complete the purchase? How strongly do you agree with: the information provided was so ambiguous that you often felt confused? To what extent do you agree: that seeking more information made completing your purchase less confusing?	Sproles and Kendall (1986)
<i>Time pressure</i>	To what extent do you agree that you had enough time to complete the purchase? How strongly do you agree that you felt like you were under time pressure while making the purchase?	De Dreu (2003)

To what extent do you agree that the presence of
a timer made you feel rushed during the
purchase?

Appendix C. Consent Form

Consent form

Purpose:

The purpose of this study is to investigate responses to shopping bots.

Equipment:

Pupil Labs eye-tracking equipment, Empatica E4 wristband, iPhone XR and MacBook Pro 16-inch 2021.

Procedure:

During this experiment, you will be asked to interact with a shopping bot. Please confirm the following. I confirm that I do not have any physical, mental or health-related reasons or problems that should preclude my participation in this study, and I also confirm that I assume all of the physical, psychological, and financial risks associated with the use of the above indicated equipment.

If you agree to participate in this experiment, you will be asked to do the following:

Interact with our Radboud shopping bot as you would normally in an online environment, while wearing the Pupil Labs eye-tracking equipment and Empatica E4 wristband. The eye-tracking equipment measures your physiological response in terms of pupil fixations and pupil dilation during the experience, and the Empatica E4 wristband logs your level of arousal. Further, your facial expressions will be recorded using the iPhone XR, and other activities will be captured using screen recording software on a MacBook Pro 16-inch 2021. The total time that is required to complete this study is approximately 15 minutes including instructions, calibrating, and debriefing.

Health notice/risk

The devices used in this experiment are not expected to cause any physical harm or other discomforting feelings. I voluntarily assume all associated risks and take full responsibility for these and any other consequences that may arise from my participation.

Confidentiality:

Your participation is voluntary, and you can leave the lab experiment at any time, without a reason or permission, even after signing this consent form. The data that is provided by you will be kept confidential and stored in a responsible and correct way. In addition, these data are only used for research purposes and not for any other secondary purposes. Personal and sensitive data will not be recorded, and the experiment is completely anonymous. The information from this research might be published in the Radboud Thesis repository. Lastly, the information and results from this project may be submitted for publication in academic journals, however, this will not include personal data and data cannot be traced back to you in any way.

Contact:

If you are interested in the results of this experiment or have any questions, please contact the research team via ..., email:...

Statement of consent:

I have read the above information; I understand this completely and I consent to participate in this experiment.

Name of Participant

Signature of Participant

Date

Thank you for your participation!

Appendix D. Script Chatbot

1. – Please interact with our Radboud shopping bot as you would normally in an online environment –
2. Hi, I am Cody, and will be your shopping assistant for the day. Before we begin, could you please provide me with your assigned participant ID?

Participants sends participant ID

3. Great, thank you! How are you doing today?
 - a. ***Answer with a positive or neutral expression:***
 - i. Great, let's continue then.
 - b. ***Answer with a negative expression:***
 - ii. That is unfortunate. I still hope you can help me out, let's continue.
 - c. ***Completely inappropriate answer: ***
 - iii. Okay, thank you for sharing, let's continue.
4. I am thankful that you are helping me today to improve my design. As a shopping bot my purpose is to enhance your shopping experience and help you complete a purchase. Let's start shopping, I will provide you with a link to a website where we can make the purchase.
5. *send link based on condition*
6. * Delete the link when they have entered the site*
 - a. ***When asking why the link was deleted*:**
 - i. The link has expired, but you can still continue on the website.
 - b. ***When asking what to do***
 - iv. Please add the purchase to your basket.
 - b. ***When asking what the timer means*:**
 - i. After three minutes, the session will expire, and the purchase cannot be completed.
7. *Wait for the error screen AND for the participant to contact shopping bot*
 - a. ***Not respond within 1 min of encountering the error*:**
 - ii. Oops something went wrong with adding the item to your basket 😞 . Please send me the error number you see on the screen so I can try to solve the error for you.
 - b. *** Asks what they should do/ask for a number for the first time*:**
 - iii. It seems you stumbled upon an error when adding the item to your basket. Please send me the error number you see on the screen so I can try to solve the error for you.
 - c. *** Asks again for clarification etc. *:**
 - iv. I cannot provide further assistance without you providing me with the error number you see. Please send me the error number you see on the screen so I can try to solve the error for you.

IF NON-CONFUSION:

Accept any code/number the participants answers

ELSE IF CONFUSION:

I am unable to use this number, please provide me with an alternative displayed error number so I can try to solve the error for you. (3X)

*** If the participant repeats the same number(s) *:** You have already provided me with this number, but I am unable to use it. Please provide me with an alternative displayed error number so I can try to solve the error for you.

8. I think I managed to sort the problem out, can you please confirm to me if the purchase has been completed by checking this link *send link*

9. *** Wait for the participant to confirm ***

a. *** If participant does not respond within 30 seconds *:**

v. Is there something wrong? I am confident that I managed to fix the error for you, can you please check this link and confirm *send link*

10. *** If participant confirms the purchase has been made *:**

a. Great to hear, it looks like we managed to resolve the issue together. To learn from this experience and help me improve for the next time, please help me in answering some questions. I would like you to reflect on your shopping experience with me. Please tell me how strongly you agree or disagree with my upcoming questions by typing in the corresponding numbers (when applicable). Let's get started.

11. Response options:

a. ***When answering in another way than a number*:**

vi. I was expecting a number, please reply to me with the respective number indicating your closest choice.

b. ***When participant answer off scale for example "8" *:**

vii. Your choice is not reflected in the answer possibilities, please reply to me with the respective number indicating your closest choice.

c. ***When participant don't understand question*:**

viii. I am afraid I am unable to provide you with further assistance. Please answer with the best of your abilities.

12. How strongly do you agree that the more information you sought, the harder it seemed to complete the purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree (5) Somewhat Agree (6) Agree (7) Strongly Agree

13. Ok, and to what extent do you agree that the provided information was clear enough to complete the purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree (5) Somewhat Agree (6) Agree (7) Strongly Agree

14. I see, and how strongly do you agree with: the information provided was so ambiguous that you often felt confused?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

15. Alright, and to what extent do you agree: that seeking more information made completing your purchase less confusing?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

16. Ok and moving on, to what extent do you agree that you had enough time to complete the purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

17. Sounds good, and how strongly do you agree that you felt like you were under time pressure while making the purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

18. Interesting... and to what extent do you agree that the presence of a timer made you feel rushed during the purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

19. Ok, thank you for sharing. To help me further, please take me through what you were thinking as you were attempting to help me resolve the purchase error we encountered.

20. I see, and what actions crossed your mind in the process of resolving this error?

21. That is insightful thank you, and would you please elaborate by telling me what was valuable in achieving the outcome of the shopping experience and how?

22. I see, so to what extent do you agree that we have accomplished what we wanted from this purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

23. Got it, and to what extent do you agree that we were effective in completing the purchase?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

24. Good, and to what extent do you agree that the way the error was resolved was useful?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

25. Moreover, how strongly do you agree that this shopping experience with me was valuable?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

26. Thank you for sharing, now in relation to how you felt, please tell me a bit about your feelings while attempting to help me resolve the purchase error.

27. Got it! And could you now please describe your overall feelings about the shopping experience outcome?

28. I see thank you for elaborating, so overall your shopping experience was ...

(1 = Extremely Displeasing, 2 = Very Displeasing, 3 = Moderately Displeasing, 4 = Slightly Displeasing, 5 = Neutral, 6 = Slightly Nice, 7 = Moderately Nice, 8 = Very Nice, 9 = Extremely Nice)

29. Clear, and in the end, how pleasant was your shopping experience?

(1 = Extremely Unpleasant, 2 = Very Unpleasant, 3 = Moderately Unpleasant, 4 = Slightly Unpleasant, 5 = Neutral, 6 = Slightly Pleasant, 7 = Moderately Pleasant, 8 = Very Pleasant, 9 = Extremely Pleasant)

30. Got it, and how agreeable or disagreeable would you say your whole shopping experience was?

(1 = Extremely Disagreeable, 2 = Very Disagreeable, 3 = Moderately Disagreeable, 4 = Slightly Disagreeable, 5 = Neutral, 6 = Slightly Agreeable, 7 = Moderately Agreeable, 8 = Very Agreeable, 9 = Extremely Agreeable)

31. Thanks for that, so this shopping experience left you feeling...

(1 = Extremely Sad, 2 = Very Sad, 3 = Moderately Sad, 4 = Slightly Sad, 5 = Neutral, 6 = Slightly Happy, 7 = Moderately Happy, 8 = Very Happy, 9 = Extremely Happy)

32. Thanks for sharing how you felt with me. So as an outcome, to what extent would you say that you are satisfied with the shopping experience?

(1) Strongly Dissatisfied (2) Dissatisfied (3) Somewhat Dissatisfied (4) Neither Dissatisfied nor Satisfied (5) Somewhat Satisfied (6) Satisfied (7) Strongly Satisfied

33. Ok, and how likely is it that you would use this shopping experience again?

(1) Highly Unlikely (2) Unlikely (3) Somewhat Unlikely (4) Neither Unlikely nor Likely (5) Somewhat Likely (6) Likely (7) Strongly Likely

34. Thanks for sharing, and how likely is it that you would tell a friend or colleague about this shopping experience?

(1) Highly Unlikely (2) Unlikely (3) Somewhat Unlikely (4) Neither Unlikely nor Likely (5) Somewhat Likely (6) Likely (7) Strongly Likely

35. Thank you and moving on, to what extent do you agree that you prefer situations in which there is some ambiguity?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree (5) Somewhat Agree (6) Agree (7) Strongly Agree

36. Alright, and how strongly would you agree that you enjoy tackling problems that are complex enough to be ambiguous?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

37. Great, and to what extent do you agree that you generally prefer novelty to familiarity?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

38. Got it, and to what extent do you agree that providing clear information is the basis for avoiding confusion?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

39. Ok, and to what extent do you agree that you want to see unambiguous information when resolving confusing issues occurring in shopping experiences?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

40. Moreover, to what extent do you agree that you enjoy tackling issues pertaining to confusion occurring from ambiguous information?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

41. Almost done, a few questions more. To what extent do you agree that I behaved like an automated assistant?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

42. Got it, and to what extent do you agree that you sensed something in your surroundings that hindered you from fully concentrating on your experience with me?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

43. Alright, and to what extent do you agree that you would have preferred chatting to a human agent as opposed to me when encountering the error we had?

(1) Strongly Disagree (2) Disagree (3) Somewhat Disagree (4) Neither Disagree nor Agree
(5) Somewhat Agree (6) Agree (7) Strongly Agree

44. Thank you so much for sharing this with me, a few last questions before you go.

Compared to how you felt starting this experience with me, how would you describe feeling now? Can you elaborate on why you feel this way?

45. Thank you for elaborating, can you please provide me with your age?

46. Ok, before you go, what is the gender you identify with the most?

(1) Male, (2) Female, (3) Other, (4) Prefer not to say.

47. Thank you for participating and being patient with me! This is the end of the experience. Please do not close any windows, stay seated, and wait till someone helps you.

Appendix E. Data Preparation

Table E1. *Univariate statistics*

	<i>N</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>Missing</i>		<i>No. of Extremes</i>	
				<i>Count</i>	<i>Percent</i>	<i>Low</i>	<i>High</i>
Age	196	25.17	8.19	1	0.5	0	18
TfAQ1	197	3.44	1.67	0	0.0	0	0
TfAQ2	197	4.73	1.39	0	0.0	6	0
TfAQ3	197	4.42	1.33	0	0.0	0	0
CAQ1	197	6.64	0.79	0	0.0	5	0
CAQ2	197	5.86	1.54	0	0.0	11	0
CAQ3	197	3.93	1.56	0	0.0	0	0
FRA	197	2143.70	1077.98	0	0.0	0	1
DRA	197	168.34	23.38	0	0.0	0	4
FNRA	197	418.82	419.40	0	0.0	0	13
DNRA	197	167.08	25.23	0	0.0	0	1
ThoughtsOQ1	197	4.32	1.08	0	0.0	2	0
ThoughtsOQ2	197	4.69	0.89	0	0.0	11	11
ThoughtsOQ3	197	3.13	0.99	0	0.0	0	2
ThoughtsCQ1	197	5.74	1.38	0	0.0	7	0
ThoughtsCQ2	196	5.70	1.63	1	0.5	16	0
ThoughtsCQ3	197	5.36	1.69	0	0.0	19	0
ThoughtsCQ4	197	5.05	1.58	0	0.0	10	0
Satisfaction	196	5.04	1.42	1	0.5	4	0
Loyalty	197	4.54	1.77	0	0.0	0	0
NPS	197	4.47	1.70	0	0.0	0	0
Realness of chatbot	196	5.76	1.21	1	0.5	7	0
Distraction	197	2.49	1.65	0	0.0	0	0
Problem-solving preference	196	3.98	1.80	1	0.5	0	0
Gender	197			0	0.0		

a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).

Appendix F. Results

Table F1. *Fornell-Larcker criterion*

	Confusion	DNRA	DRA	FNRA	FRA	Loyalty	Satisfaction	Thoughts Closed
Confusion	1							
DNRA	0.04	1						
DRA	-0.085	0.532	1					
FNRA	0.01	0.152	0.026	1				
FRA	0.19	-0.008	0.237	-0.232	1			
Loyalty	-0.225	0.05	0.092	-0.003	-0.111	0.829		
Satisfaction	-0.315	0.03	0.087	-0.015	-0.118	0.737	1	
Thoughts Closed	-0.392	0.057	0.045	-0.003	-0.138	0.707	0.75	0.785

Table F2. *Heterotrait-monotrait ratio (HTMT)*

	Confusion	DNRA	DRA	FNRA	FRA	Loyalty	Satisfaction	Thoughts Closed
Confusion								
DNRA	0.040							
DRA	0.085	0.532						
FNRA	0.010	0.152	0.026					
FRA	0.190	0.008	0.237	0.232				
Loyalty	0.274	0.07	0.094	0.026	0.152			
Satisfaction	0.315	0.03	0.087	0.015	0.118	0.854		
Thoughts Closed	0.451	0.063	0.063	0.110	0.157	0.909	0.829	

Table F3. *Path coefficients and significance*

Hypothesis	Path	β	t-value	p-value	Support
H1a:	Customer Confusion \rightarrow satisfaction	-0.395	3.115	<0.001***	Yes
H1b:	Customer Confusion \rightarrow Loyalty	-0.228	1.606	0.054	No
H2:	Customer Confusion \rightarrow Total Fixation RA	0.318	2.062	0.020*	Yes
	Customer Confusion \rightarrow Total Fixation NON-RA	-0.038	0.274	0.392	No
	Customer Confusion \rightarrow Average Duration RA	-0.172	1.073	0.142	No
	Customer Confusion \rightarrow Average Duration NON-RA	0.174	1.070	0.142	No
H3:	Customer Confusion \rightarrow Thoughts	-0.551	4.453	<0.001***	Yes

H4a:	Customer Confusion → Total Fixation RA → Satisfaction	-0.006	0.294	0.385	No
	Customer Confusion → Total Fixation NON-RA → Satisfaction	-0.001	0.112	0.455	No
H4b:	Customer Confusion → Average Duration RA → Satisfaction	-0.014	0.758	0.224	No
	Customer Confusion → Average Duration NON-RA → Satisfaction	-0.016	0.780	0.224	No
H4b:	Customer Confusion → Total fixation RA → Loyalty	-0.005	0.256	0.399	No
	Customer Confusion → Total fixation NON-RA → Loyalty	-0.001	0.088	0.465	No
H5a:	Customer Confusion → Average duration RA → Loyalty	-0.016	0.819	0.206	No
	Customer Confusion → Average duration NON-RA → Loyalty	-0.010	0.587	0.279	No
H5a:	Customer Confusion → Thoughts → Satisfaction	-0.378	4.189	<0.001***	Yes
H5b:	Customer Confusion → Thoughts → Loyalty	-0.404	4.215	<0.001***	Yes
H6a:	Customer Confusion → Total Fixation RA → Thoughts → Satisfaction	-0.001	0.070	0.472	No
	Customer Confusion → Total Fixation NON-RA → Thoughts → Satisfaction	-0.002	0.187	0.426	No
H6a:	Customer Confusion → Average Duration RA → Thoughts → Satisfaction	0.002	0.121	0.452	No
	Customer Confusion → Average Duration NON-RA → Thoughts → Satisfaction	0.002	0.187	0.426	No
H6b:	Customer Confusion → Total fixation RA → Thoughts → Loyalty	-0.001	0.069	0.427	No
	Customer Confusion → Total fixation NON-RA → Thoughts → Loyalty	-0.002	0.196	0.422	No
H6b:	Customer Confusion → Average duration RA → Thoughts → Loyalty	0.002	0.121	0.452	No
	Customer Confusion → Average duration NON-RA → Thoughts → Loyalty	0.002	0.188	0.426	No

*p < 0.05, **p < 0.01, ***p < 0.001

Table F4. Path coefficient, t-value, and p-value of control variables

Path	β	t-value	p-value
Age -> Duration Non-Relevant	-0.124	1.800	0.036*
Age -> Duration Relevant	-0.223	3.310	<0.001***
Age -> Fixations Non-Relevant	0.037	0.387	0.349

Age -> Fixations Relevant	0.015	0.180	0.428
Age -> Loyalty	-0.071	1.267	0.103
Age -> Satisfaction	-0.071	1.241	0.107
Age -> Thoughts Closed	-0.145	2.064	0.020*
Realness of Chatbot -> Duration Non-Relevant	0.095	1.420	0.078
Realness of Chatbot -> Duration Relevant	0.064	1.009	0.157
Realness of Chatbot -> Fixations Non-Relevant	-0.098	0.98	0.164
Realness of Chatbot -> Fixations Relevant	-0.013	0.158	0.437
Realness of Chatbot -> Loyalty	0.064	1.177	0.120
Realness of Chatbot -> Satisfaction	0.085	1.411	0.079
Realness of Chatbot -> Thoughts Closed	0.021	0.308	0.379
Distraction -> Duration Non-Relevant	0.145	1.610	0.054
Distraction -> Duration Relevant	0.092	1.155	0.124
Distraction -> Fixations Non-Relevant	0.096	1.150	0.125
Distraction -> Fixations Relevant	0.003	0.039	0.484
Distraction -> Loyalty	0.037	0.698	0.243
Distraction -> Satisfaction	0.078	1.578	0.057
Distraction -> Thoughts Closed	0.018	0.276	0.391
Problem-Solving Preference -> Duration Non-Relevant	-0.063	0.783	0.217
Problem-Solving Preference -> Duration Relevant	0.087	1.095	0.137
Problem-Solving Preference -> Fixations Non-Relevant	0.109	1.877	0.030*
Problem-Solving Preference -> Fixations Relevant	0.114	1.504	0.066
Problem-Solving Preference -> Loyalty	-0.064	1.094	0.137
Problem-Solving Preference -> Satisfaction	-0.106	1.811	0.035*
Problem-Solving Preference -> Thoughts Closed	-0.352	5.337	<0.001***
Confusion Avoidance -> Duration Non-Relevant	0.009	0.132	0.448
Confusion Avoidance -> Duration Relevant	0.032	0.432	0.333

Confusion Avoidance -> Fixations Non-Relevant	0.066	1.346	0.089
Confusion Avoidance -> Fixations Relevant	0.017	0.219	0.413
Confusion Avoidance -> Loyalty	-0.032	0.664	0.253
Confusion Avoidance -> Satisfaction	-0.065	1.371	0.085
Confusion Avoidance -> Thoughts Closed	0.002	0.034	0.487
Gender -> Duration Non-Relevant	0.319	2.385	0.009**
Gender -> Duration Relevant	0.247	1.799	0.036*
Gender -> Fixations Non-Relevant	-0.252	1.787	0.037*
Gender -> Fixations Relevant	0.114	0.755	0.225
Gender -> Loyalty	0.106	1.120	0.131
Gender -> Satisfaction	0.021	0.207	0.418
Gender -> Thoughts Closed	0.055	0.408	0.342
Tolerance for Ambiguity -> Duration Non-Relevant	0.024	0.255	0.399
Tolerance for Ambiguity -> Duration Relevant	-0.032	0.284	0.388
Tolerance for Ambiguity -> Fixations Non-Relevant	-0.116	1.458	0.072
Tolerance for Ambiguity -> Fixations Relevant	0.038	0.405	0.343
Tolerance for Ambiguity -> Loyalty	-0.003	0.054	0.478
Tolerance for Ambiguity -> Satisfaction	-0.002	0.042	0.483
Tolerance for Ambiguity -> Thoughts Closed	0.057	0.752	0.226

*p < 0.05, **p < 0.01, ***p < 0.001