Understanding the impact of customer confusion on the customer experience with a chatbot

Master Thesis

Master Thesis of Business Administration

Innovation and entrepreneurship

Name student: **Bas Waters**
Student number: **S1087341**

Supervisor: **K. Sidaoui**
Second examiner : **R.H.W. Wetzels**

Radboud University, Faculty of management
MSc Innovation and Entrepreneurship
Nijmegen 30-06-2022

Radboud University

Source picture front page UFM (n.d.)

Abstract

Self-service technologies (SSTs) like chatbots have become unthinkable in today's society. 80% Of businesses utilize chatbots to engage with their customers, as they have proven to be efficient, cost effective, and convenient. Nevertheless, chatbots can also induce confusion in customers, which is known to affect customer experience (CX) outcomes like satisfaction and loyalty greatly. Businesses often do not know during a CX what kind of thoughts and sensations their chatbot trigger during customer confusion and how this affects business outcomes. This thesis utilizes an Wizard of Oz type of participant study to research the mediating effect of sensation and thoughts on the relationship between customer confusion and CX outcomes satisfaction and loyalty. Participants (n = 200) engage with a chatbot during a shopping experience in which confusion is induced. During this, sensation is assessed by measuring gaze behavior with eye tracking and thoughts on the service are estimated by a questionnaire. PLS-SEM analyses results indicate a mediating effect of positive thoughts on the relationship between customer confusion and satisfaction and loyalty, but not of increased gaze behavior. The relationship between customer confusion and gaze behavior was partially significant, meaning that increased gaze behavior is an indicator of customer confusion. In addition, confusion was significantly related to negative thoughts on the automated service. These confirmative as well as new findings indicate that businesses can benefit from evaluating their chatbots by assessing customers thoughts on the CX to increase customer loyalty and satisfaction.

Inhoudsopgave

1. Introduction

The exploitation of self-service technologies (SSTs) has increased greatly in customer service and has become unimaginable in today's society (Curran et al., 2003). Such services include ATMs, self-check-in and outs in hotels and chatbots. Using these SSTs in marketing enables companies to help customers without direct contact with an employee (Meuter et al., 2000). SSTs have proven advantageous for companies because of their accuracy, cost-effectiveness, and efficiency (Zhu et al., 2007). For example, it was estimated that chatbot adoption in business operations could result in yearly savings of more than $8 billion by 2022 in retail, ecommerce, banking, and healthcare (JuniperResearch, 2017). Chatbots are used for the convenience of both end-users and businesses because of their ease of use, usefulness, and low cost (Przegalinska et al., 2019). Approximately 80% of companies already use chatbots to interact with customers and resolve their issues at all times (Ramasamy, 2019). As mentioned, the SSTs have also proven to be meaningful for customers. For instance, it was found that customers experience an increased 'perceived control', meaning that customers think they have more influence over the service process and outcomes of SSTs (Zhu et al., 2007).

Besides the advantages of SSTs for both customers and businesses, the rapid adoption of these technologies has also led to increased customer confusion. Customer confusion can be defined as the disorientation and loss of understanding when customers are confronted with unknown situations, uncertainties, errors or information that counters their preconceptions (D'Mello & Graesser, 2014; Pachman et al., 2016a). According to one study, 89% of businesses observed staff or customer confusion when introducing new products or services (Easingwood, 1986). While traditionally, services have been given by an employee, customers using SSTs have to get used to this new interaction, which can be challenging for some individuals (Meuter et al., 2005). Companies might be persuaded to automate their customer support completely to reduce operational costs. However, companies often incorrectly assume that their customers are already familiar with using these new SSTs (Collier et al., 2015). Because of this, unwanted confusing situations can occur. For example, a chatbot might confuse the customer rather than quickly solve the issue. The chatbot could also depreciate value by reacting in a way that conflicts with the company's image or identity (Canhoto & Clear, 2020). When the chatbot fails to provide the

customer with the knowledge they need, customers might think that the chatbot is worthless. On the other hand, when the SSTs provide an overload of information, customers may get stressed and frustrated and make unrealistic judgments, which could confuse them even more (V. Mitchell & Papavassiliou, 1999).

SSTs that cause confusion in costumers can lead to poor business outcomes (Canhoto & Clear, 2020). Research shows that negative emotions caused by confusion during the customer journey affect overall customer satisfaction more than positive ones (Babin & Darden, 1996). Moreover, when there is an overload of information and amibiguity confusion is induced, the satisfaction of customers is also affected (Walsh & Mitchell, 2010b). Besides customer satisfaction, it is also known that confused customers are less likely to reuse the automated service and to return to a company (Meuter et al., 2005) In other terms, confusion negatively influences the customers loyalty (V. Mitchell & Papavassiliou, 1999).

Confusion caused by SSTs can harm the customer experience (CX) in multiple ways. CX is characterized by five dimensions during the interaction between a customer and a company: thought, feeling, sensation, activity, and relation (De Keyser et al., 2020; Schmitt, 1999). Companies tend to know a lot about their customers' purchasing patterns, incomes, and other traits which are used to predict the CX. However, they know little about customers' actual thoughts, feelings, and states of mind during the interaction with their SSTs (Meyer & Schwager, 2007). This study will focus on the subdimensions sensation and thoughts and their role in the relationship between confusion and customer experience outcomes satisfaction and loyalty. Sensation involves the sensory system of vision. Confusion has already been associated with increased gaze behavior (Salminen et al., 2018). Therefore, gaze behavior could be an interesting predictor of customer satisfaction and loyalty. In addition, customers cognitively process information obtained during their interaction with a SST, for example through their vision. These thoughts say something about their utilitarian values on the service (Babin et al., 1994), which are expected to predict customer satisfaction and loyalty.

1.1 Research objective and question

SSTs have been widely adopted over recent years because of all their advantages for businesses and customers. As explained, these technologies can also increase customer confusion, ultimately affecting the overall CX in terms of customer satisfaction and loyalty. The CX can be assessed by the sub-dimensions thoughts and sensations.

This research aims to understand the relationship between customer confusion and CX outcomes satisfaction and loyalty and how this relationship is mediated by gaze behavior (sensations) and thoughts. The following research question will be answered:

"*What is the effect of customer confusion on the thoughts and sensations during the customer experience, and therefore ultimately on customer satisfaction and loyalty?"*

1.2 Relevance

Customer confusion affecting the overall CX is important to marketers because confused customers are less likely to make rational purchasing choices, choose items with the highest quality or value for money and enjoy the customer journey (Walsh & Mitchell, 2010a). Furthermore, customer confusion has been linked to various economic effects for companies, such as word of mouth, cognitive dissonance, decision delay, dissatisfaction, reduced trust, and reduced loyalty (Walsh & Mitchell, 2010). Marketing managers often experience difficulty detecting what aspects of their technologies trigger confusion and frustration. Considering everything, it is important to investigate what triggers customer confusion while using SSTs. This thesis will provide more insight into customers' thoughts and sensations while experiencing the actual negative or positive CX. This information is relatively new, as CX is often evaluated after these thoughts and sensations have already occurred. Therefore, the manner in which customer satisfaction was harmed cannot be estimated properly.

Marketing managers can use this valuable information to improve their SSTs, such as chatbots, so the technology can intervene when confusion is detected during the CX. These alterations can improve CX, and companies can benefit from the described advantages that come with it, such as increased customer satisfaction and loyalty. Increased satisfaction of customers can lead to more positive word of mouth, which will lead to a higher reach of possible new costumers while this form of advertisement is cost free for the company. Logically, new customer's as well

as increased loyalty of old customers will increase growth of the company as well as its turnover. Moreover, an improved shopping experience with SSTs will reduce the need for other employees to resolve issues that were not fixed by the SST. Also, a more efficient shopping experience can result in reduced lead time in the company's operational processes, making the SST available for more customers in a shorter time period. All of these situations will lead to reduced costs for companies.

1.3 Thesis outline

The main concepts of the research question have been mentioned in the introduction. This thesis will further elaborate on these concepts in the literature review, in which all of the hypotheses are mentioned and reasoned as well. Then, the methods of the conducted participant study and of the (statistical) analyses are mentioned in the methodology section. The results of these analyses are discussed in the results section after which these results are interpreted and put into context in the discussion section.

2. Literature review

The structure of this chapter is as follows: At first, the role of self-service technologies (SSTs) will be discussed, as the rapid development of these technologies is one of the main reasons this research is conducted. Secondly, the concept of customer confusion that can be caused by these SSTs will be explained and what kind of impact this confusion has on CX. Thirdly, the concept of CX will be defined with the main focus on the CX dimensions of 'thoughts' and 'sensations'. At last, customer satisfaction and customer loyalty will be explained. The conceptual model will illustrate all the corresponding hypotheses in the last paragraph.

2.1 Self-service technologies (SSTs)

The development of technology involved in the delivery of services has complicated the service industry, which was traditionally driven by interpersonal interactions. As service providers add touchpoints, consumers have additional technology-based service delivery alternatives without direct interaction with service business staff (Curran et al., 2003). These services include chatbots, self-check-in and out at hotels, and ATMs (Meuter et al., 2000). SSTs promise better cost efficiency, service quality, and customer acquisition over in-person services (Zhu et al., 2007). As a result of using SSTs, clients feel they have greater "perceived control" meaning that they feel they can affect the outcome of the service process themselves (Zhu et al., 2007). Technological advancements are becoming increasingly important in customer–firm interactions (Meuter et al., 2000). For example, self-increasing service relevance is mentioned by Parasuraman (2000) as a fundamental change like service. However, little is known about how these technology solutions affect CX and customer behaviour (Meuter et al., 2000).

In the context of SSTs, the primary focus of this thesis will be on the technology of chatbots because of several reasons. First of all, chatbots were implemented by businesses because of their advantages over human services. Chatbots are virtual assistants that are claimed to assist a lot of customers individually and effectively as they are more flexible than human agents and often 24/7 available (Schmitt, 2020). In addition, because chatbots are easy to use, useful, and low in costs, this SST is used for the convenience of both end-users and businesses (Przegalinska et al., 2019). By 2022, adopting chatbots in company operations could save the retail, e-commerce, banking, and

healthcare sectors more than $8 billion annually (JuniperResearch, 2017). Besides the advantages of the chatbot, the chatbot has become unthinkable in today's businesses. Chatbots have been considered a significant technological trend since 2016 (Baier et al., 2018). They are utilised by 80% of businesses to engage with customers (Ramasamy, 2019). The majority of businesses that use chatbots employ them for customer support (95%), followed by marketing (60%) (Mindbrowser, 2020).

2.2 Customer Confusion

While SSTs certainly have their benefits, their rapid adoption can also contribute to a rise in customer confusion. Cognitive disequilibrium, which refers to a condition an individual experiences during learning when barriers to the usual learning process flow are met, is the cause of confusion (D'Mello & Graesser, 2014). These barriers include uncertainty, mistakes, unusual information, or simply new information contradicting an individual's knowledge. In this state, a person is sometimes confused about what to do next (D'mello & Graesser, 2013). Due to information overload and confusion, consumers may become stressed, frustrated and make poor decisions (V. Mitchell & Papavassiliou, 1999).

Mitchell et al. (2005) define three categories of confusion resulting from brand similarity, information overload, and misleading or ambiguous information. In terms of the definitions, confusion is an aware state of mind that can occur either before or after a purchase and consists of a cognitive as well as a behavioural component (Mitchell et al., 2005).

Brand similarity confusion can be defined as *'a lack of understanding and potential alteration of a consumer's choice or an incorrect brand evaluation caused by the perceived physical similarity of products or services' (Mitchell et al., 2005).*

Confusion caused by information overload is described as *'a lack of understanding caused by the consumer being confronted with an above information rich environment that cannot be processed in the time available to understand, and be confident in fully, the purchase environment (Mitchell et al., 2005)'.*

Ambiguity perplexity is defined as *'a lack of understanding during which consumers are forced to re-evaluate and revise current beliefs or assumptions about products or the purchasing environment' (Mitchell et al., 2005).*

An example of ambiguity confusion during the interaction with a chatbot: The chatbot asks the customer for a code but is not specific enough. Subsequently, the customer sees multiple options that could be the code on an interface and does not know which code to provide to the chatbot.

Businesses may be convinced to fully automate their customer support in an effort to cut costs. However, confusion can be triggered by multiple situations during the interaction of customers with chatbots. For instance, the chatbot could depreciate value by reacting in a way that conflicts with the company image or identity, this could be against a customer's expectation and could therefore cause confusion (Canhoto & Clear, 2020).

Importantly, confusion has been linked to several negative outcomes (Mitchell et al., 2005). Firstly, customer confusion negatively affects the customer experience outcomes satisfaction and loyalty. For example, increased customer confusion leads to lower repeat sales and more returned goods (Lu & Gursoy, 2015). Also, confusion in customers has been linked to a negative brand image, negative word of mouth, cognitive dissonance, delayed decision-making, and distrust (Walsh & Mitchell, 2010).

2.3 Customer Experience

The customer experience (CX) has become a crucial aspect of modern marketing (Trivedi, 2019) and can be highly affected by customer confusion. CX refers to the interaction between a customer and a company along five dimensions: thought, emotion, sensation, activity, and relationship (De Keyser et al., 2020; Schmitt, 1999). Typical 'features-and-benefit' marketing of organisations seek to build an experience for their customers (Schmitt, 1999). This interaction may impact customers' thoughts (Trivedi, 2019) and give businesses a competitive edge (Clarke & Kinghorn, 2018). De Keyser et al. (2015) also mentioned that a positive CX would provide any company with a competitive advantage, which was supported by marketing academics and practitioners. A customer's expectations and the stimulus they experience as a result of interacting with a business and its products are compared to the various touchpoints or times of contact. (LaSalle & Britton, 2003; Shaw & Ivens, 2002).

De Keyser et al. (2020) break down the complexity of CX management into its basic parts: touchpoints, context, and attributes (TCQ). The first building block, touchpoints, is essential for experience development since there can be no CX without real or perceived contact (Hoffman &

Novak, 2018). Context generally refers to the conditional state that defines the resources a person may, directly and indirectly, use at a given moment (Bettencourt et al., 2014; Chandler & Vargo, 2011). The qualities include a group of distinguishing characteristics that represent the type of responses and reactions customers have to their encounters with the brand or company (De Keyser et al., 2015; Keiningham et al., 2020). The TCQ language clarifies CX and facilitates the identification of elements that influence CX. The TCQ terminology will enable three significant contributions to CX research and practice. At first, the TCQ terminology reduces conceptual uncertainty and confusion among CX academics and practitioners by identifying and listing all CX components. Secondly, the TCQ terminology enables us to evaluate the present status of academic CX work and identify knowledge gaps. Thirdly, the TCQ terminology makes CX actionable and helps companies to achieve their CX goals via business practice (De Keyser et al., 2020). In order to construct a solid relationship, managers need to identify the touchpoints that have an effect on CX and establish what kinds of contextual data should be gathered. (De Keyser et al., 2020).

Companies know a lot about their consumers' experience shopping habits, income levels, and other characteristics that distinguish them, but little about the thoughts, feelings, and states of mind their encounters with SSTs produce, such as chatbots (Meyer & Schwager, 2007). Schmitt (1999) was the researcher who developed a detailed framework for experiential marketing. It has two parts: Strategic Experience Modules (SEMs) and ExPros (short for experience producers). Schmitt (1999) believed that people are emotional and rational creatures who want to experience pleasurable experiences. Based on this logic, he constructed five "strategic experiential modules" (SEMs) to summarise the entire experience. These SEMs are characterised by the following characteristics: *"sensory experiences (SENSE), affective experiences (FEEL), creative cognitive experiences (THINK), physical experiences, behaviors, and lifestyles (ACT), and social-identity experiences associated with a reference group or culture (RELATE)"* (Schmitt, 1999). The definitions of each dimension are explained in Table 1. According to Schmitt (1999), the final aim of marketers is to combine the five encounters to create a comprehensive experience.

Table 1

Customer experience (CX) dimensions sense and think described by Schmitt, 1999 and their definition.

Dimensions	Definition
Sense	SENSE marketing speaks to the senses to generate sensory experiences through sight, hearing, touch, taste, and scent.
Think	THINK marketing speaks to intelligence, intending to generate cognitive, problem-solving encounters that involve consumers creatively. THINK appeals involve clients' convergent and divergent thinking through astonishment, mystery, and provocation.

The sub-dimensions of sensation and thoughts were used in this thesis to conceptualize CX as a single concept. The other dimensions of Schmitt fall outside the scope of this research. To explain this decision, sensations refer to the customer's physical feelings and perceptions while thoughts refer to the customer's mental evaluation of the service or product they have received. Eye tracking is able to detect visual processing, mind wandering, cognitive load, visual attention and even confusion (Lee et al., 2021; Salminen et al., 2018) Thus, vision could potentially indicate a lot about a CX. For example, in literature, increased gaze behavior has already been linked to confusion (Meißner & Oll, 2019). In addition, when customers are confused by SSTs, they may form negative thoughts about the service or product, which can have a lasting impact on their perceptions. Thoughts can be measured by utilitarian values, these values have become important considerations during assessing CX (Evelina et al., 2020). These findings indicate that sensation and thoughts are predicted to say a lot about CX and thus also on CX outcomes loyalty and satisfaction.

2.3.1 Sensation

Krishna (2012) defines sensory experience as a marketing technique that uses sensory input to influence customer evaluation, decision-making, and behavior. It aims to create new experiences by appealing to the customer's senses of sight, sound, taste, touch, and smell to attract more customers (Krishna, 2012; Lindstrom, 2005a). Schmitt (1999) says that the five senses give consumers a fuller picture of a brand and make it easier for them to connect with it, which makes the product more valuable. According to Lindstrom (2005b), sensory experience is critical in

developing an emotional connection with a product. It is the basis of experiential marketing's five pillars, and the approach will likely leave a long-lasting impression on customers.

Certain sight- and sound-focused sensory marketing methods have lost their novelty. Nevertheless, according to a survey by Lindstrom (2005a), 37% of participants believed that sight is an essential sensation, followed by smell (23%), sound (20%), taste (15%), and touch (5%). On the other hand, Lindstrom (2005a) argues that the value of the five senses is comparable. Consequently, organisations should not only target the sight and hearing senses, but also provide possibilities to trigger the other three. Lindstrom (2005b) indicates that all sensations must be adequately stimulated to connect goods or brands with customer emotions effectively, therefore encouraging consumer loyalty. Applying the sensory experience to produce pleasurable experiences is a crucial business problem (H.-T. Chen & Lin, 2018).

As mentioned earlier, eyesight is an important sensation according to customers. Granka et al. (2008) emphasize that most eye-tracking user research is focused on studying cognitive information processing: what users' thoughts are and what they are paying attention to. The main difficulty is distinguishing positive or desired cognitive states (for example, curiosity) from negative ones (for example, confusion). The investigation of general problem-solving and circumstances involving confusion has benefited from using the method of eye tracking (Pachman et al., 2016a). In the eye-tracking method, fixations are measured to confirm confusion. A fixation refers to a constant eye position lasting 100–600 ms while focusing on a certain gaze point (Salminen et al., 2018). Fixations reveal in which aspect more information may be required, and processing may occur by capturing the direction of a user's attention (Granka et al., 2008). Shorter, quicker eye movements between fixations are known as saccades (Salminen et al., 2018). Information from eye tracking can confirm confusion because gaze behaviour describes users' gaze patterns regarding these amount of fixations and the duration of fixations (Salminen et al., 2018). Thus, vision in the form of gaze behavior can be used as a measure to assess CX.

2.3.2 Thoughts

Thoughts in marketing refer to the intelligence to generate cognitive, problem-solving interactions that engage customers creatively (Schmitt, 1999). Astonishment, mystery, and provocation trigger convergent and divergent thinking in customers (Schmitt, 1999). Marketing campaigns, as well as the way products are designed, can provoke certain thoughts in customers. Subsequently, this

affects how customers experience a certain service, as well as whether they are eventually satisfied and are likely to reuse the service again. It appears that creative thinking comprises two main aspects caused by two ways of thinking: the creation of novelty (through divergent thinking) and the evaluation of the novelty produced (through convergent thinking) (Cropley, 2006). When services provoke convergent thoughts in customers, they are more prone to be able to have highly effective decision-making, which will result in a more efficient shopping experience. In addition, a positive or negative customer experience can affect the utilitarian values of customers, which depends on whether the specific consumption needs that led to the shopping trip were achieved (Babin et al., 1994). Usually, this means that a product was bought in a planned and efficient way. Both customer satisfaction and client loyalty should be connected to utilitarian values and studied (Babin et al., 1994; Jones et al., 2006). When shoppers perceive utility value in their purchases, they are likely to complete the purchasing "task" of making a purchase (Babin et al., 1994; Jones et al., 2006). Thus, the higher the utilitarian value, the more likely customers are satisfied and will reuse the service. Therefore, thoughts in the form of utilitarian values can be assessed to determine CX.

2.4 Customer Satisfaction and Loyalty

Customer satisfaction is a beneficial result of minimal customer confusion and a positive CX and has been found to be essential to a company's success and long-term competitiveness (J.-S. Chen et al., 2021). Oliver (1993) defined customer satisfaction as the consumers' judgment that a product or service feature, or the product itself, provided a better experience for customers of consumption-related fulfilment, including under or over-fulfilment (Oliver, 2014). Transaction-specific and cumulative or overall satisfaction are the two most commonly accepted methods for studying customer satisfaction (Srivastava & Kaul, 2014). According to the transaction-specific approach, customer satisfaction is defined as the consumer's emotional reaction to the most recent experience with an organisation (Oliver, 1993). Customers' cumulative or overall evaluation of a company's service performance and the level of satisfaction with particular goods and different aspects of the business are reflected in the overall satisfaction (Srivastava & Kaul, 2014). According to Molla and Loicker (Molla & Licker, 2001), customer satisfaction is a reaction and feeling linked to the customer's experience.

Companies often think customers know how to use SSTs, however, when this is not the case, customer confusion can be triggered, and therefore a negative or harmed CX can occur (Collier et al., 2015). Customer confusion has been tied to economic impacts for companies like word of mouth, cognitive dissonance, choice delay, dissatisfaction, lower trust, and loyalty (Walsh & Mitchell, 2010). In addition, if a chatbot does not meet the customer's expectations. The chatbot could decrease customer satisfaction, discouraging service adoption (Ashfaq et al., 2020).

According to researchers, loyalty is a major outcome of customer satisfaction (Cheng & Jiang, 2020). Customer loyalty refers to a person's attachment to a product, company, retailer, service provider, or other organisation based on positive attitudes and reactions like repeated purchases (Zakaria et al., 2014). Customer loyalty is crucial for a company to keep its existing customers, as they can profit greatly from loyal customers (Zakaria et al., 2014). Customers loyal to a brand are less likely to shop around for the best price, which lowers the company's overall marketing costs and boosts its probability of selling (Rowley, 2005). Customers' enjoyment and fun in chatbots may lead to an improved brand image, more potential purchases, and greater brand recognition (Cheng & Jiang, 2020), which enhances customer loyalty. In conclusion, customer satisfaction and loyalty is highly beneficial for companies and SSTs like chatbots should not negatively affect these CX outcomes.

2.5 Time urgency

The literature hypothesizes that the stress caused by perceived crowding increases the negative effect of using SSTs, such as self-checkout (Gelbrich & Sattler, 2014). Perceived time pressure refers to the extent to which a person perceives himself or herself to be time-constrained concerning daily responsibilities (Alreck et al., 2009). Although Creating a service "on-site" usually demands that customers complete the service when it is their turn, it might become important for SST usage (Gelbrich & Sattler, 2014). According to Maule et al. (2000), people under time pressure frequently experience stress, which may amplify the negative effects of technological anxiety. Also, when customers experience confusion due to time pressure, they may change their gaze behavior by lowering the duration and number of fixations (Pieters & Warlop, 1999). However, the effect of time urgency on confusion and altered gaze behavior is not yet fully known, whereas another pattern could be observed in gaze behavior when there is

presence of time urgency. This information could be valuable for companies to take into account and apply this to their specific kind of services.

2.6 Conceptual model and hypotheses

2.6.1 Direct effects of confusion on customer satisfaction and loyalty

Information overload that causes customers to invest more effort in decision-making might lead to dissatisfaction because too much information from which not everything is relevant is provided to the customer (Turnbull et al., 2000). When overload and ambiguity confusion emerges, customers have to commit more time to process the information, which could negatively influence customer satisfaction (Walsh & Mitchell, 2010b). Subsequently, when customers experience confusing situations, this could also negatively affect their thoughts on the usage of the service, which indicates that customers are less likely to return to the company (Vincent-Wayne & Vassilios, 1999). Taking the above into consideration, the following two hypotheses are drafted:

> **H1a.** Increased customer confusion has a negative relationship with customer satisfaction in a chatbot encounter.
>
> **H1b:** Increased customer confusion has a negative relationship with customer loyalty in a chatbot encounter.

2.6.2 The direct effect of confusion on sensations while moderated by time urgency

Schmitt (1999) mentioned that one of the five CX dimensions is a sensory experience, and eyesight seems to be the most essential sensation (2005a). Granka et al. (2008) emphasize that the primary challenge is to distinguish the cognitive state of a human mind, which is perceived as desirable (e.g., curiosity), from those considered negative (for example, confusion). According to several articles, researchers suggest that increased cognitive load (confusion) increases the total amount of fixations and that the average fixation duration is longer (Marshall, 2002; Salminen et al., 2018). Besides the direct effect of confusion on gaze behaviour, this research will also investigate the relationship between customer confusion and sensations moderated by time urgency, as this moderating effect has not been studied before. However, it is likely that time urgency creates stress in customers and that this affects their gaze behaviour. Because of this, the following two hypotheses are suggested:

H1C: Time urgency negatively moderates the relationship between customer confusion and increased gaze behaviour in a chatbot self-service.

H2: Customer confusion is positively related to increased gaze behaviour in a chatbot self-service.

2.6.3 The indirect effect of CX sensations on customer satisfaction and loyalty

To build on the information provided above, according to Jimenez-Bareto (2021), chatbot self-services can improve the sensory experience by understanding client requests and providing information that stimulates customers' senses. On the contrary, chatbot interactions induce confusion due to misleading answers or technical errors, adversely negatively impacting information processing (Turnbull et al., 2000). When the whole self-service stimulus is too complex for the customer, this can create overload confusion (Sharma et al., 2023; Turnbull et al., 2000). This indicates that confusion and satisfaction or loyalty are related in a way mediated by gaze behaviour. This leads to the following hypotheses:

H3a: Increased gaze behaviour negatively mediates the relationship between customer confusion and customer satisfaction in a chatbot self-service.

H3b: Increased gaze behaviour negatively mediates the relationship between customer confusion and customer loyalty in a chatbot self-service.

2.6.4 The indirect effect of CX Thoughts on customer satisfaction and loyalty

According to research, utilitarian values gained during shopping experiences have a direct effect on customer satisfaction scores (Babin et al., 1994; Machleit et al., 2000). Nevertheless, the situational context of the experience is important. Customer satisfaction in fast-food restaurants is more greatly affected by utilitarian factors, focused on completing a goal or task, than by hedonistic factors, focused on pleasure and fun (Ryu et al., 2010). So, for example, in the context of a fast-food restaurant convenience and speed of service should be more important evaluation criteria. Brandtzaeg and Folstad (2017) indicate that customers' use of chatbots can be tied to utilitarian aims consisting of efficiency, productivity, and convenience. Furthermore, according to Jones et al. (2006), customer loyalty and customer satisfaction should be connected to utilitarian value. Utilitarian values are expected to positively influence customer satisfaction, but this relationship between utilitarian value and customer satisfaction is stronger compared to utilitarian value and customer loyalty. In this thesis, it is essential to identify what kind of thoughts arise during CX when customers are confused. Therefore, the following hypotheses are formulated:

H4: Customer confusion is negatively related to positive thoughts in a chatbot self-service.

H5a: Positive thoughts positively mediate the relationship between customer confusion and customer satisfaction in a chatbot self-service.

H5b: Positive thoughts positively mediate the relationship between customer confusion and customer satisfaction in a chatbot self-service.

2.6.5 The effect of customer confusion via CX sensations and CX Thoughts on customer satisfaction and customer loyalty

In this research, we are interested in the process of how confused customers undergo the CX dimensions (sensations and thoughts) of Schmitt (1999) and what kind of effect this has on customer experience outcomes satisfaction or loyalty. The ability of customers to efficiently process information cognitively may be affected by several forms of confusion, such as similarity, overload, and ambiguity (V.-W. Mitchell et al., 2005). As a result, this might cause people to process information less cognitively (Meißner & Oll, 2019). Furthermore, Lindstrom (2005b) indicates that all senses must be adequately stimulated to effectively link products or brands with customer emotions, fostering consumer loyalty. According to Chen and Linn (2018), applying the sensory experience to produce pleasurable experiences is a crucial business problem. Moreover, Jones et al. (2006) suggested that there should be a correlation between customer loyalty and customer satisfaction, and utilitarian value. Utilitarian values have been hypothesised to have a positive effect on customer satisfaction. However, research has indicated that the relationship between utilitarian values and customer satisfaction is more robust than the relationship between utilitarian values and customer loyalty (Jones et al., 2006). Consequently, this thesis will clarify the indirect relationship between customer confusion and customer satisfaction and loyalty. It is anticipated that increased consumer confusion may generate more gaze behaviour (sensations), negatively affecting cognitive load (positive thoughts) and ultimately negatively influencing customer satisfaction and loyalty. Taking everything into consideration, the following hypothese are formulated:

H6: An increase in gaze behaviour negatively influences the positive thoughts in a chatbot self-service.

H7a: Negative thoughts and increased gaze behaviour negatively mediate the relationship between customer confusion and customer satisfaction in a chatbot self-service.

H7b: Negative thoughts and increased gaze behaviour negatively mediate the relationship between customer confusion and customer loyalty in a chatbot self-service.

2.6.6 Control variables

Four control variables are included in this study: Age, gender, tolerance of ambiguity, and confusion avoidance (see Figure 1). These control variables have been incorporated because of their potential impact on the study results. The first control variable is age. In research on chatbots and other types of SSTs, age is frequently used as a control variable (Adam et al., 2021; Nguyen et al., 2022; Rietz et al., 2019). Given that higher age is associated with a decreased confidence in technological systems, it is estimated that older people have less trust and confidence in SSTs (Adam et al., 2021).

The second control variable is gender. The findings of Eliot and Hall (2005) mention differences between males and females regarding readiness and willingness to use self-service technologies. The willingness and technology-readiness for SST is stronger in males than females (Elliott & Hall, 2005). Females prefer to interact with an employee of an organization instead of with a SST.

The third control variable is tolerance for ambiguity. Generally, the degree to which a person is comfortable with uncertainty, unpredictability, competing demands, and conflicting directions is known as their tolerance for ambiguity (McLain, 2009). If people are uncomfortable in an unknown situation, this can reduce their intention to use a new SST. Less desire to use technology may lead to preconceptions about SSTs, affecting how people deal with and remember a chatbot service (Sidaoui et al., 2020).

The last control variable is confusion avoidance. Customers learn to act in certain ways to protect themselves from too many triggers, also from confusion (Schweizer et al., 2006). If they learn themselves this new set of behaviors, they can respond differently to a SST in the future. Hence, this variable should be taken into account.

2.6.7 Conceptual model

The proposed relations and hypotheses are all visualized in the conceptual model (Figure 1) underneath.

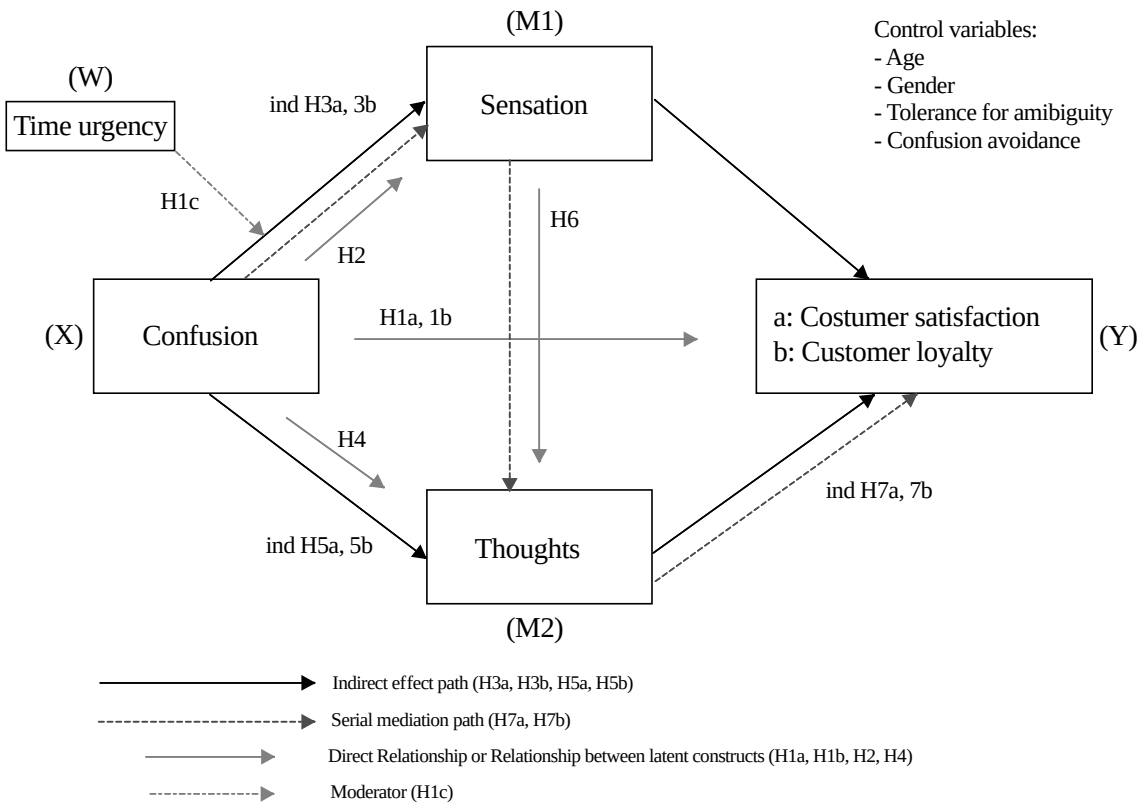


Figure 1: Conceptual model of this study

3. Methodology

The data collection process and analysis will be covered in the following chapter. The research strategy, sample size & sampling method, software applications, analysis technique, operationalisation, ethical considerations, and limitations will be described in this chapter.

3.1 Research Strategy

3.1.1 Scenario

In this research, participants interacted with the shopping bot Cody. After participants agreed to participate in the experiment, they were introduced to the Pupil Labs eye-tracker (see Figure 2) and the chatbot 'Cody' displayed on the monitor in front of them (see Figure 3). Each participant obtained an unique ID code that they had to enter in the Whatsapp Business interface to start the conversation with Cody. After the first general questions, the participants received a hyperlink that they had to open to start the shopping experience with Elite Forté.

The first page of the site showed a context-free product or service named 'Premium Comfort X', which the participants had to add to the shopping basket to continue the customer journey. However, the shopping experience was immediately interrupted by an error message showing an error number or several other relevant numbers. The display of the error message depended on which confusion condition the participants were in.
Additionally, besides the error message, it was also possible that participants saw a timer. This also depended on which time urgency condition the participants were in. There were four different conditions, which will be further explained below (paragraph 3.2.1.). During the error message page, the participants had to work together with Cody by providing a specific code displayed on the site to solve the problem. When the correct code was sent to Cody, Cody sent a link back which shows that the product was successfully ordered, and that the participant can wait

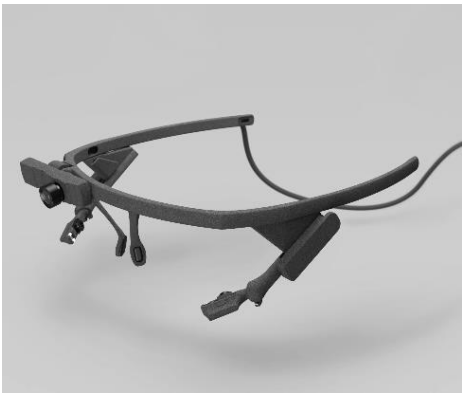


Figure 1: Pupil Labs core eye-tracker (Pupil Labs, n.d.)

for their product to arrive. Subsequently, Cody took a questionnaire from the participant to determine his/her customer experience with Cody and the webshop Elité Forté.

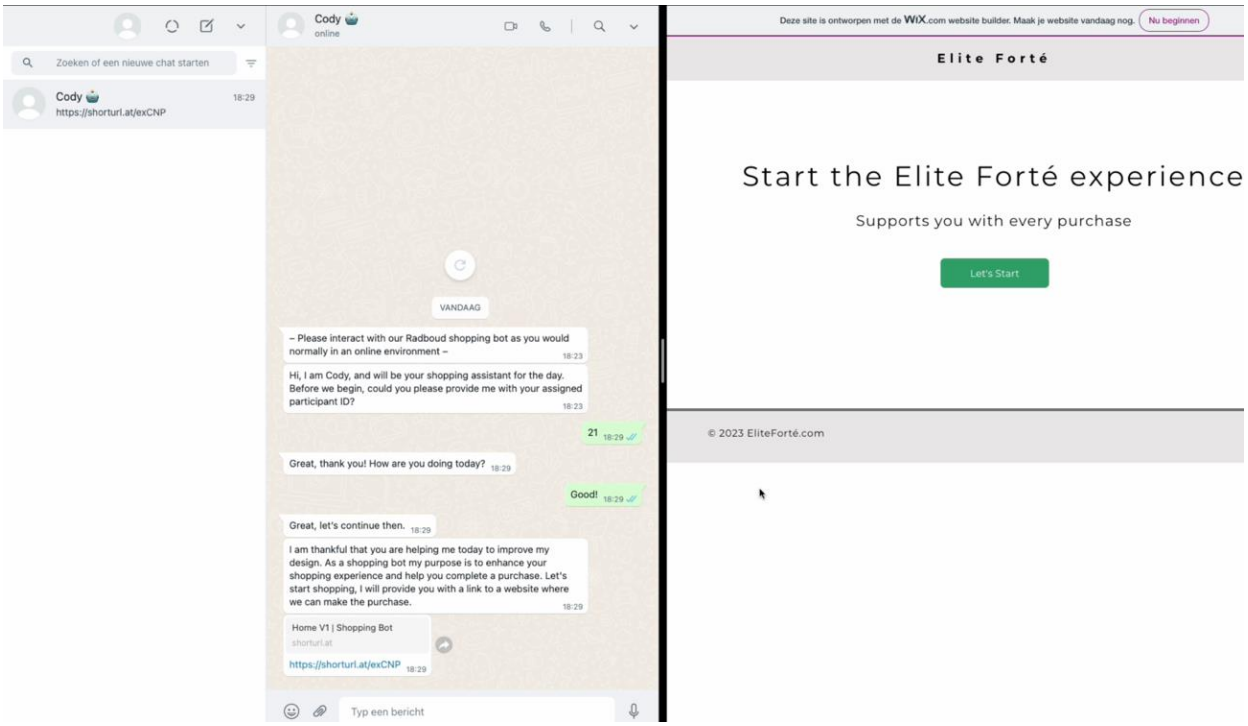


Figure 2: Screen that the participants see at the beginning of the experiment with the shopping bot 'Cody' and the webshop of Elite Forté

3.2 Experimental design

Before participants were recruited for the experiments, a pre-test was performed with some individuals to check if the manipulation check was likely to induce confusion and time urgency. When this information was confirmed, participants were recruited. The Pre-test will be discussed in paragraph 3.2.2.

Participants completed and filled in their informed consent before entering the experimental setting, ‘the lab’. Then, the participants received the instruction for the experiment and the unique ID code required to start the conversation with the shopping bot ‘Cody’. The consent form is shown in Appendix 1. To start the experiment, two or three members of the thesis circle set up everything to start the experiment, such as changing the seating height, installing the Empatica wristband, changing the positioning of the eye-tracking equipment, calibrating the eye-tracking equipment,

and starting all the recording equipment in the lab. Eye tracking was performed during the whole experiment.

During this experiment an Wizard of Oz technique was applied. The chatbot (the wizard) was in reality one of the researchers who started to chat with the participant when the participant typed in their ID number in the Whatsapp Business interface (Riek, 2012). However, the participant was told that he/she would talk to the newly developed shopping bot 'Cody'.

At the beginning of the experiment, the participants received a hyperlink which started the experiment. The participants had to click the 'start' button to start the customer experience (see Appendix 2, figure 2.1). The second page showed a shopping page in which the participant had to add the context-free product or service 'Premium Comfort X' (see Appendix 2, figure 2.2). An error message would appear when the participants clicked the 'add to shopping basket' button. Subsequently, the participants had to collaborate with Cody to solve the technical problem. The error was solved when the participants entered the right error number that Cody asked for. However, the display of the error message depended on what type of condition the participants were in. There were two confusion conditions and two time urgency conditions: In the 'No' confusion condition (see Appendix 2, figure 2.3), a simple error message was shown with only one obvious error number. In the 'Yes' confusion condition (see Appendix 2, figure 2.4), an error message was displayed which included multiple numbers with no obvious error number, such as article number, order number, etc. In the 'No' time urgency condition (see Appendix 2, figure 2.3), a 3 minute timer was displayed on the screen. In the 'Yes' time urgency condition (see Appendix 2, figure 2.5), there was no 3 minute timer displayed on the screen. When the participants solved the problem with Cody, the shopping bot would sent a new fixed page. This page ensured that the product or service 'premium Comfort X' was successfully purchased (see Appendix 2, figure 2.7).

After the shopping experience, the chatbot asked questions about confusion, time urgency, utilitarian value, and other relevant control variables. This information was used to evaluate the thoughts and sensations of the participants during the confusion and confirm whether confusion took place. In addition, it was assessed how satisfied participants were with the service and whether they would use the service again in the future.

After all questions had been answered, a member of the thesis circle group removed the eye-tracking equipment and shortly debriefed with the participant. Finally, a thesis circle group member thanked the participant for participating in the experiment.

3.2.1 Experimental conditions

There are four conditions in which the participants were tested: with confusion (yes or no) and with time urgency (yes or no). Between these conditions there were differences in what the participants would see on the display as described above. During the experiments, the participants were unaware of the experimental condition in which he or she was allocated to. The conditions were divided over the participants at random by letting participants drew a number after they signed the consent form. The first condition serves as a control group as in this condition no confusion is induced as well as no time urgency present. In the second condition, confusion is induced when no time urgency is present. In the third condition, confusion is not induced while time urgency is present. In the last condition, confusion and time urgency are induced.

3.2.2 Pre-test

A pre-test was conducted to ensure that the two conditions of confusion and time urgency were correctly experienced (Field, 2013). An one-way ANOVA was used to analyze the data obtained during the pre-test (Field, 2013). In total, 20 participants were recruited for the pre-test. From these participants, ten participants experienced the confused condition, and ten participants experienced the time urgency condition. During the pre-test experiment, participants had to answer nine questions based on a 7-point Likert scale about the experienced confusion and time urgency. For an overview of the pre-test results, see Appendix 3. According to the descriptive statistics, the participants perceived the confusion condition as confusing (mean = 5,08) and the time urgency condition to be a time pressure situation (mean = 4,87) (see Appendix 3). The findings of the pre-test indicate that all confusion items' mean differences between groups (confusion: yes and no) are statistically significant for three items ($p < .001$), except for confusion item 2 ($F(1,18) = 2.654$; $p = 0.121$). According to the test's results, mean differences between groups for all time urgency items are statistically significant for two items $P < 0.05$. The first item of time urgency is insignificant ($F(1,18) = 3.534$; $p = 0.076$). Therefore, there are no significant mean differences between these two groups.

3.3 Sample size and participant recruitment

In this research, non-probability convenience sampling was applied to reach individuals around the campus of the Radboud University. This sampling method takes the population being studied into consideration. A downside is that this method is limited due to the fact that only probability

sampling generates representative results. Consequently, external validity cannot be encouraged according to Field (2013). Nevertheless, on the Radboud University campus the use of non-probability convenience sampling is highly applicable and cost-effective, and data can be easily obtained.

Sample sizes of 15-20 observations per variable are optimal for regression analysis (Field, 2013). If the ratio is less than 5:1, the possibility exists that there is no disproportionate fitting of the variate to the sample (Field, 2013). The proposed conceptual model consists of 11 variables, which leads to a sample size that ranges between 165 and 220 participants with a minimum of at least 55 participants. Eventually, 200 participants were recruited in this study. From these 200 participants, eye tracking data of two participants was unusable due to technical problems with data export and of 1 participant the recording was not complete. These three participants were removed the analysis (Hair et al., 2018). Thus, data of 197 participants were utilized in the data analysis.

3.4 Analyses method and software applications

In this study six software applications were utilized. The first program was SmartPLS, which was used for statistical analyses with Partial Least Square Structured equation modelling (PLS-SEM) and confirmatory factor analyses (CFA) using bootstrapped data were performed with SmartPLS. Hair et al. (2018) employed SmartPLS to assess serial mediation in the model, which was also done in this study. SmartPLS allows a variety of intricate routes, such as running two mediators and one moderator in one PLS model simultaneously (Hair et al., 2018). The second program, SPSS Statistics, was applied for data analysis as well. SPSS was used for data cleaning, reverse coding, independent sample t-tests and ANOVA analyses (IBM SPSS Statistics, 2020). The third program was Whatsapp Business, which was used during the experiment as a chatbot-similar interface to create the Wizard of Oz simulation. Whatsapp business is known for allowing chatbots on its platform, and it has a professional interface, including timestamps. These timestamps are important for data export (Whatsapp, n.d.). The fourth program was Wix.com, which was used during the experiment as well. The website helped with creating a realistic simulation of a web shop for the participant. This made it easier for the participants to imagine and facilitate the experiment's different conditions (Wix.com, n.d.). The fifth program was Pupil Labs, which recorded the eye tracking data during the experiment measured with the Pupil Labs glasses (PupilLabs, n.d.). The last program, Phyton TextBlob library, was used to analyze the open questions to assess the

subdimension thoughts. This software determines and scores whether texts were negative, neutral or positive and can be assessed on the internet for free (Textblob, n.d.).

3.5 Construct measurement

Regression analysis is the best technique for models with continuous mediating variables and categorical independent variables (Iacobucci, 2012). The aim of this research is to determine a serial indirect impact for which moderated serial mediation would be the most suitable option (Demming et al., 2017).

3.5.1 Serial moderated mediation

The conceptual model (Figure 1) implies a causal chain connecting the mediators in serial mediation and provides a more comprehensive analysis of the complexity of multiple X variables, which impacts Y compared with a single mediation model (Montoya & Hayes, 2017). The effectiveness of a serial mediation model is evaluated in this thesis. As stated in the chapter 'literature review', it is anticipated that there will be a causative relationship between the suggested mediators. Also, it is expected that time urgency moderates the relationship between customer confusion and sensations. The control variables will be introduced as covariates to the moderated serial mediation model to control for every kind of common relationship between model variables due to external causes (Hayers, 2012).

3.5.2 Operationalization

This research makes use of modified measures that were already used and described in literature (See Appendix 4 , Table 4.1). Eye-tracking technology was used to measure gaze behaviour in order to assess the construct of CX sensations. The evaluation consisted of three items adapted from Meißner and Oll's (2019), Pachman et al. (2016), and Salimin et al. (2018). The construct of thoughts was divided into two dimensions: open questions and utilitarian values. The open-ended question items assessed the participants' in-depth thinking, allowing them to express their thoughts freely and in their own words (Fossey et al., 2002; Leiva et al., 2006). Three items assessed with the open questions have not been previously studied in research and the significance of these elements was determined via sentiment analysis. On the other hand, the utilitarian value construct consisted of four items with a 7-point Likert scale that have been adjusted and drawn from the work of Babin et al. (1994).

The CX outcomes, customer satisfaction and loyalty, were assessed based on the chatbot's interactions. A semantic scale was utilized to measure customer satisfaction. While the specific source of this scale is unknown, it bears a similarity to other scale questions commonly used in customer satisfaction literature. This approach allowed customers to express their satisfaction levels, ranging from 1, "strongly dissatisfied", to 7, "strongly satisfied". Regarding customer loyalty, two items from a prior study by Reichheld et al. (2003) were incorporated and evaluated on a seven-point Likert scale. This semantic scale spans from 1, "highly unlikely", to 7 "strongly likely".

The measures of confusion and time urgency are important components of the experiment. These scales were important in determining whether the intended confusion and time urgency were effectively induced. However, it is important to note that these items were only relevant for the conduction of a manipulation check and did not contribute to the analysis within SmartPLS.

The measurement of confusion entailed four items, which had been adapted from Sproles and Kendall's (1986) study. Utilizing a 7-point Likert scale, these items were specifically designed to measure the level of confusion experienced. Notably, to ensure suitability of these questions in our own experimental scenario with the chatbot, the items from Sproles and Kendall's (1986) study were thoughtfully modified and rewritten to reflect a conversational tone that would resonate with users.

Three items were included to measure the level of time urgency based on the research conducted by De Dreu (2003). These items utilize a 7-point Likert scale to measure the intensity of time urgency experienced. To ensure their suitability within the chatbot context, these items were carefully adapted and rephrased to reflect a conversational tone as well.

This research also incorporated four control variables: age, gender, confusion avoidance, and tolerance for ambiguity. Age and gender are demographic variables that are not directly related to the literature under investigation, therefore the chatbot asked for this information during the questionnaire. Tolerance for ambiguity was assessed using three 7-point Likert scale items, adapted from Mclain's (2009) research and rephrased as narrative questions instead of statements. However, the three items for confusion avoidance, based on relevant literature by Schweizer et al. (2006), were adapted to align with the specific requirements of the research and were also measured using a self-made 7-point Likert scale that is similar to the other questions in

the study. The comprehensive list of items corresponding to each construct can be found in Appendix 4.

3.6 Data preparation

The data preparation involved applying sentiment analysis on the open question items (Gujjar & Kumar, 2021), assessing the validity, dealing with missing values, missing value analysis and evaluating its reliability (Field, 2013). The thoughts open items were based on a sentiment analysis scale that ranges from -1 (negative) to +1 (Positive). However, in this research only 7-point scales were used. Therefore, the scales of the sentiment analysis were converted to a 7-point scale.

Furthermore, data was gathered for different conditions, including confusion (yes and no confusion) and time urgency (yes and no time urgency). These datasets were combined into a single SPSS data matrix for further analysis. The experimental study comprised 200 participants, but three participants had to be excluded from the dataset due to insufficient eye-tracking data caused by technical challenges. Additionally, certain items, namely MCconfusion Q2, MCconfusion Q4, MCtimeurgency Q1, and CAQ3 were subjected to reverse coding as they were originally formulated in the opposite direction.

An examination was conducted to assess missing values in the data. This step of the analysis helped to determine whether the missing values occur completely at random (MCAR) or if they followed a random pattern (MAR). In the case of MAR, it suggests that the absence of certain data points could be attributed to a specific factor within the dataset. This scenario can potentially impact the reliability of the dataset, as there might be a connection between the missing data and the respondent's decision not to answer a particular question (Field, 2013). The study of missing values revealed that the percentage of missing data for each item was significantly lower than the cutoff point of 5% suggested by Hair et al. (2018). Little's MCAR test revealed that all of the missing values in the data set were fully random missing values: x^2 (185) = 139,194; p=.062. No further issues with the missing values appeared.

Moreover, the data preparation process involved classifying the confusion and time urgency conditions using dummy variables to compare the groups' confusion and time urgency (Hair et al., 2018).

At last, regarding the eye-tracker data, fixations and average fixation duration within different areas of interest (AOIs) were initially divided into AOI surface 1, AOI surface 2, and non-relevant areas to prepare the eye-tracking data for analysis. Therefore, the fixations and average duration of fixations across these surfaces must be combined in this study. This consolidation enabled a thorough examination of the eye-tracking data, taking into account all areas of interest collectively.

3.7 Research Ethics

This thesis shall uphold the values of honesty, scrupulousness, transparency, independence, and responsibility (KNAW, 2020).

Honesty in keeping good records of the research process and not ignoring different points of view or arguments to argue against, making unproven claims, making up data, or modifying data or information to make the results look better.

Scrupulous use of scientific and academic procedures. Carefully constructing the thesis, reporting appropriately, and referencing sources following the approved American Psychology Association (APA) referencing style.

The independence of research shall be maintained. Rejecting non-scientific or non-scholarly influences on procedure, data, or assessment determinations. Independence shall be maintained throughout the thesis by not engaging other parties in the research process.

Transparency will be guaranteed. To begin, study participants are assured of confidentiality and anonymity. Before conducting the questionnaire with the chatbot, participants must agree to the terms of data sharing. Secondly, the method by which the data is processed, utilised and evaluated using statistical tests to reach conclusions will be shown step by step. Finally, the methodology part will be transparent by explaining the step-by-step process.

Consideration of accountability was taken into account when demonstrating scientific and practical relevance. The study is relevant since SSTs, like chatbots, are projected to be immensely valuable. Furthermore, a comprehensive literature analysis has produced a strong theoretical foundation for this thesis.

All the data collected from WhatsApp Business during the experiment is securely stored in temporary meta datacenters across Europe. These meta data centers are in Denmark, Ireland, and

Sweden (Meta, n.d.). The selection of these geographical locations adheres to stringent data protection standards and regulations in Europe.

4. Results

To analyze the data, both SPSS and SmartPLS software were utilized. In short, SPSS was conducted to gain insights into the descriptive statistics and assess the normality of the data. Furthermore, SmartPLS 4 was employed to conduct a comprehensive investigation of the measurement model and structural model, encompassing assessments of reliability, validity, hypotheses, and assumptions. SmartPLS 4 is a software package specifically designed for performing partial least squares structural equation modeling (PLS-SEM) analyses on the dataset (Hair et al., 2018). Furthermore, a bootstrapping technique was employed to obtain robust estimates, specifying 5000 subsamples of bootstrapping procedure (Hair et al., 2018).

4.1 manipulation check

To check the manipulation of confusion and perceived time urgency in the experiment, data of 197 participants were analyzed with an independent t-test and ANOVA. The findings from these tests revealed significant mean differences between the groups (confusion: yes and no) for all the items related to confusion ($p < .001$). Regarding the mean differences between the groups based on time urgency (yes and no), two out of the three items showed significant differences ($p < .05$). From these three items, item 1 did not yield a significant result ($p = .532$).

4.2 Evaluation of the measurement model

The measurement model was evaluated with the following steps: model fit, inter-reliability, convergent validity and discriminant validity. All thresholds for this analysis are shown in Table 4. Additionally, all the steps of the analysis are in appendix 5. The SRMR measure showed that the theoretical model fit was favourable (Dijkstra & Henseler, 2015). The SRMR score measured how strongly the empirical correlation matrix deviated from the model's implied correlation matrix in addition to evaluating the model fit. Hu and Bentler (1999) found that values below 0.05 denote a good fit and that a value below 0.08 indicates an acceptable fit. A sufficient fit was indicated by the results' SRMR score of 0.039 (see Appendix 5, Table 5.9).

Three different estimates were utilized to assess the construct reliability: Dijkstra-Henseler's rho (ρA), Jöreskog's rho (ρc), and Cronbach's alpha (α) (Cronbach, 1951; Dijkstra & Henseler, 2015; Werts et al., 1978). After deleting the following items: ThoughtsOQ2, ThoughtsOQ3, TfaQ2, TfaQ3. All thresholds of the earlier mentioned construct reliability were

not violated anymore. All these measures surpassed the predefined threshold of 0.7, indicating satisfactory construct reliability based on the established measurement criterion. The obtained scores higher than 0.7 are presented in Table 4. To evaluate the constructs, the average variance extracted (AVE) was computed (Fornell & Larcker, 1981). The AVE values exceeding 0.5 were considered sufficient, following the suggestion by Hair et al. (2018). All the steps of evaluating the construct reliability and validity are shown in Appendix 5 from Table 5.9 until Table 5.23.

Discriminant validity was assessed using the Heterotrait-Monotrait ratio of correlations (HTMT). This measure helps to determine the distinctiveness of items within the construct. According to Henseler et al. (2015), lower HTMT values indicate greater distinctiveness, with a suggested threshold below 0.9 and preferably below 0.85. The correlation between construct CX thoughts, CX sensations, and satisfaction did not exceed the < .850 HTMT threshold (see Appendix 5, Table 5.25). However, loyalty exceeded this threshold, so item NPS was deleted. After the deletion, the HTMT threshold was not violated; therefore, the discriminant validity was sufficient (see Appendix 5, Table 5.26). At last, According to the Fornell-Larcker criterion, the AVE measurements should be higher than the squared correlations for all other model items (Fornell & Larcker, 1981), see Appendix 5, Table 5.24). All the steps of evaluating the discriminant validity are in shown in Appendix 5 from Table 5.24 until Table 5.26.

Table 4: Measurement model Thresholds

Measurement technique	**Measurement model Threshold**	**Results**
Model fit	SRMR < 0.08	SRMR value: 0.039
Inter consistency reliability	- Dijkstr- Henseler's ρA > 0.7 - Jöreskog's rho ρc > 0.7 - Cronbach's α > 0.7	- pA: All items > .817 - pc: All items > .865 - α: All items > .791
Convergent validity	AVE > 0.5	All items > .616
Discriminant validity	HTMT < 0.85	All items < .829
Fornell-Lacker criterion	Fornell-Lacker Item > 0.5	All items have no item correlation > 0.5

4.3 Structural model

The evaluation of the structural model involved several steps, including the assessment of collinearity among predictor constructs, examination of effect size, determination of the coefficient of determination, and the significance of specific indirect effects and total effects (Hair et al., 2018; Henseler, Hubona, & Ray, 2016).

4.3.1 Collinearity among predictor constructs

The degree of collinearity was assessed using the variance inflation factor (VIF) among predictor constructs, as suggested by Hair et al. (2018). VIF values below 5 were considered acceptable as a threshold for collinearity (Hair et al., 2018). When conducting the collinearity measure, everything was within the <5 VIF threshold. See Appendix 5, Table 5.27.

4.3.2. Coefficient of determination

The coefficient of determination, which indicates the model's predictive power, was calculated. The R2 was chosen as an appropriate measure for the coefficient of determination (Hair et al., 2018). Furthermore, values above 0.75 are considered substantial, those above 0.50 are considered moderate, and those above 0.25 are considered weak (Hair et al., 2018). A threshold of >0.10 or lower is acceptable, depending on the research context (Hair et al., 2018). In our study we found the following: Total fixations relevant AOI, total fixation non-relevant area, fixation duration non-relevant area and fixation duration relevant AOI, ThoughtsCQ and ThoughtsOQ were considered extremely weak (<.25). The constructs satisfaction and loyalty were considered as moderate (>.50). At last, SmartPLS did not measure the predictive power of all the control variables. An overview of the coefficient determinations are shown in Appendix 5, table Table 5.30 and Table 5.31.

4.3.3 Effect size

The f2 is seen as a good measure for the coefficient of determination (Hair et al., 2018). The R2 was calculated using a threshold greater than 0.02. (Hair et al., 2018). Expanding on this threshold, effect sizes were classified as follows: greater than 0.35 indicates large effects, greater than 0.15 indicates moderate effects, greater than 0.02 indicates small effects and less than 0.02 indicates no effect (Hair et al., 2018). As a result, effect sizes range from 0.000 to 0.784. The total fixation non-relevant area (0,000) has the smallest effect size observed (see Appendix 5, Table 5.29). Thoughts closed question to loyalty (0,784) has the largest effect size observed (see Appendix 5, Table 5.29)

4.3.4 Significance of specific indirect effects and total effects

The significance of the Beta coefficient provides useful insight into whether the hypothesized effects match the structural model results. The significance of these coefficients allows us to determine whether the data supports the expected relationships. In figure 4 all the R-square scores are encountered in the conceptual model. However, sensations exist of multiple measurements, which made it not possible to include the Beta coefficient and the path significance. Therefore, all the results per hypotheses are shown in Appendix 6, Table 6.1.

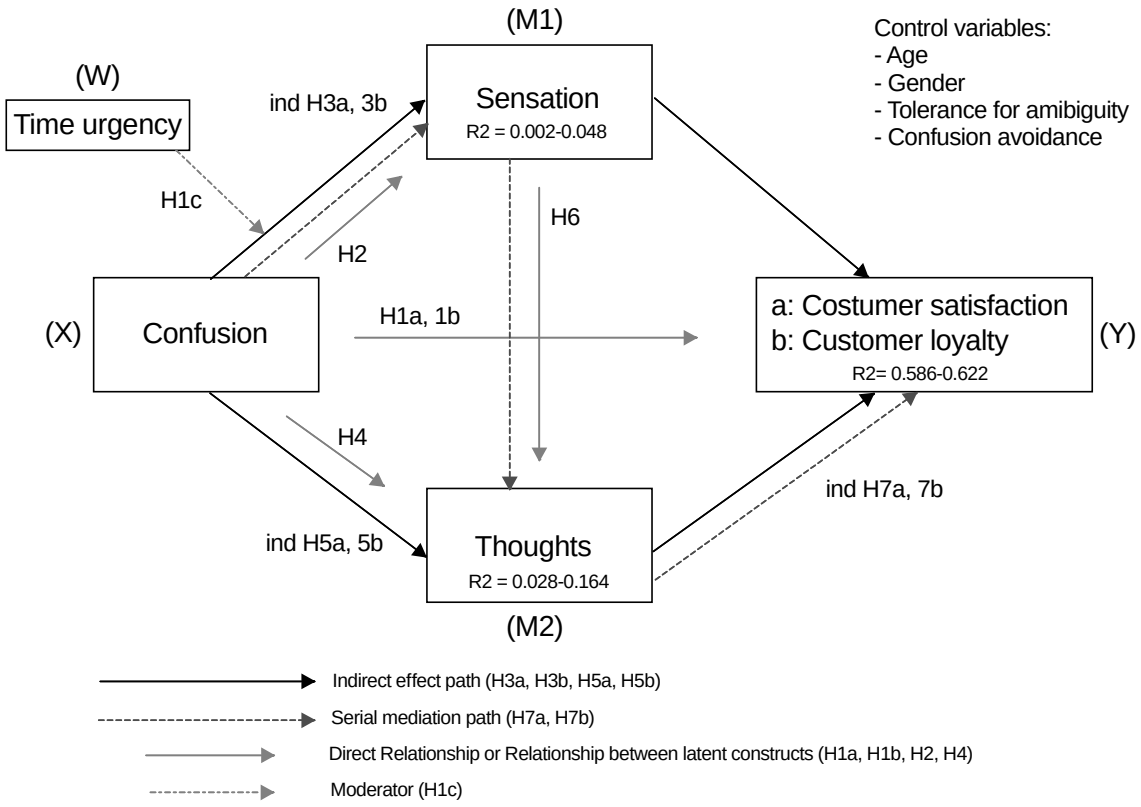


Figure 3: Model with R-square values

The presence of confusion and time urgency in the confusion ('yes') and time urgency ('yes') was validated with the manipulation check. Customer confusion while utilizing the chatbot self-service has a significantly negative relationship with customer satisfaction ($p < 0.001$) as well as with customer loyalty ($p < 0.05$) (H1a, 1b). Sensation in the form of gaze behaviour does not mediate the relationship between customer confusion and customer satisfaction, and customer loyalty (all p values > 0.05) (H3a, 3b). On the contrary, negative thoughts on the service obtained in utilitarian

values negatively mediate the relationship between confusion and customer satisfaction ($p < 0.001$) and loyalty ($p < 0.001$) and negative thoughts on the service during customer confusion obtained via open questions only negatively mediate customer satisfaction ($p < 0.05$) but not loyalty ($p > 0.05$) (H5a, 5b). Customer confusion is partially positively related to sensations, as confusion is significantly related to 2 of the 4 variables measured for gaze behaviour ($p < 0.05$) (H2). The relationship between confusion and sensation is partially moderated by time urgency because time urgency moderates the relationship between confusion and 2 of the 4 variables measured for gaze behaviour significantly ($p < 0.05$) (H1c). Confusion has a significant negative relationship with positive thoughts for both results of thoughts obtained from utilitarian values ($p < 0.001$) as well as from open questions ($p < 0.05$) (H4). There is no relationship between sensation in terms of gaze behaviour and results on positive thought obtained from utilitarian values and open questions (all p values > 0.05) (H6). The sensations in terms of gaze behaviour and thoughts in that specific order do not mediate the relationship between customer confusion and customer satisfaction or loyalty (all p values > 0.05) (H7a, 7b). All control variables do not have a significant relationship with customer loyalty or with customer satisfaction (all p values: > 0.05)

5. Discussion

The results mentioned in the results section will be discussed in this chapter. The key findings will be examined and interpreted. Also, the implications of the findings for the theory and managerial domains will be examined in light of prior research. After this, the research's limitations will be evaluated and recommendations for further studies will be made.

5.1 Key findings

This study gave insight into the mediating effect of sensation and thoughts on the relationship between customer confusion and customer loyalty and confusion and satisfaction. From these two dimensions, only thoughts appeared to mediate these relationships.

5.2 Theoretical contributions

In literature, it was already known that customer confusion negatively affects customer satisfaction and loyalty (Turnbull et al., 2000; Walsh & Mitchell, 2010a). Results from this study were in accordance with this, as there was a significantly negative relationship found between confusion and satisfaction and confusion and loyalty. The customer experience subdimensions sensations and thoughts have been described by Schmitt (1999) but were never researched yet on their involvement in confusion and its link to customer experience outcomes. In this study, we researched the mediating effect of thoughts and sensations of individuals on the relationship between how well they were confused during an automated service and how they score their customer loyalty and satisfaction.

Confusion had already been linked to increased gaze behaviour in literature (Salminen et al., 2018), which was also indicated by this study as participants had more and longer fixations on non-relevant areas when they were confused during an automated service. Unexpectedly, we found that sensation did not mediate between confusion and customer loyalty and confusion and satisfaction. In addition, we found that sensation explains very little variance of the observed data in the structural model as the R2 is very low. On the contrary, the subdimension thoughts had a relatively higher R2 and appeared to mediate the relationship between confusion and loyalty and confusion and satisfaction. In literature, utilitarian values have been linked to customer satisfaction rather than customer loyalty (Jones et al., 2006). In this study, we found that thoughts measured in utilitarian values mediated the relationship between confusion and satisfaction as well as confusion

and loyalty. This emphasizes the importance of customers thoughts during an automated service and how this can influence customer experience outcomes.

Schmitt has described the subdimensions of sensations and thoughts but does not elaborate on the relationship between these two dimensions. Some studies have looked into the link between gaze behaviour and cognitive load (Perkhofer & Lehner, 2019) or cognitive styles (Raptis et al., 2017; Steichen et al., 2020). However, our study sheds light on the unexplored link between gaze behaviour and utilitarian values to say something about the relationship between sensations and thoughts. We found that there was no significant relationship between these two subdimensions. Not entirely unexpected, we found that the effect of confusion on customer satisfaction and loyalty was not significant through gaze behaviour and thoughts.

5.3 Managerial contributions

Besides the theoretical contributions of our findings, our results also implicate some managerial contributions that can be beneficial for businesses on short terms. Firstly, our results imply that managers should research what customers thoughts are on and what their utilitarian value is of the interaction with a chatbot. For instance, whether they think the interaction was valuable. This can assist them with the design of a chatbot that fits with the company's value proposition. For example, is the company solving its customers' specific problems with their current chatbot? When this is not the case and this remains unfound, the chatbot continues to cause a lot of confusion in customers. Therefore, customers thoughts should be evaluated as our results indicate that low utilitarian values decrease loyalty as well as satisfaction in confused customers.

Secondly, although gaze behavior was not a mediator of the relationship between confusion and satisfaction and confusion and loyalty, there was a partially significant relationship found between confusion and gaze behaviour. This relationship was also found in literature. Thus, measuring gaze behavior of customers with eye tracking during a chatbot encounter could indicate when confusion is triggered and when the chatbot should intervene or aid more help to the customer. This could result in lead time reduction. Nevertheless, implementation of this could be difficult because of privacy related issues. But, managers could still test their chatbots before implementation with eye tracking to discover points of improvement in their chatbots to try to avoid customer confusion during the CX at all. In addition to this, managers could use this eye tracking method for testing to continuously update and train their chatbots to provide accurate, reliable and relevant information to their customers at all times.

Lastly, the relationship between confusion and decreased satisfaction and loyalty was already implied in literature, but also again confirmed in our results. This, again, highlights the importance of reducing or eliminating confusion in customers during a CX. Businesses should re-evaluate whether their chatbot causes some kind of confusion, as this greatly affects important business outcomes. They can not only assess customers thoughts, as this was implied to be important by our findings, but can also focus on other dimensions mentioned by Schmitt (1999) if this has their preference.

5.4 Limitations

This study provides new insight into the theoretical contributions as well as the managerial contribution. However, this study also has some limitations. The main limitation is that the Pupil Labs Core eye-tracker had some hardware issues regarding the right eye-camera which was therefore removed from the glasses. Unfortunately, this right camera could not be replaced and measurements were performed with only the left camera. This led to some inaccurate measurements on the left side of the interface, as these are normally captured by the right camera, and difficulties with calibrating the left eye-tracker camera. It was still possible to perform the eye tracking, however, the effect of this technical issue is not fully understood. For instance, whether this affected the measurements for gaze behaviour and if this is a potential reason why increased gaze behavior was not found to mediate the effect of confusion on satisfaction and loyalty. In addition, measurements with the eye-tracker were sometimes hindered by the fact that people wore eye make-up, glasses, or contact lenses. Lastly, the experiment created a situation in which participants had to purchase a context-free product or service. During the debrief with participants after the experiment, some participants mentioned that they sometimes felt like they were in a fictional or non-real situation because the information on what they were buying was missing which is not the case during a normal shopping experience.

5.5 Future research

From this study, serval future research directions can be proposed. To start with, the sensation ‘sight’ was explored in this study, as it was found that sight was thought to be the most essential sensation by customers. Results indicate there was no mediating effect of gaze behaviour on the relationship between confusion and customer experience outcomes loyalty and satisfaction. However, it should be noted that the other sensations, such as sound or smell, could have a

mediating effect and that other sensations should not be forgotten. Thus, future research could look into other sensations than sight. In addition to this, as our eye-tracking was potentially less reliable because of an impaired right camera, future research could include gaze behavior again to test its mediating effect on confusion and CX outcomes. Secondly, if a similar study is performed, it should be taken into account that when participants are put in a situation in which there is time urgency, the participant's sensation is affected. Results indicate that time urgency reduces total fixations and that the fixation duration is shorter when participants are confused compared to conditions with no time urgency. Thirdly, it is recommended to continue this research, especially for the CX dimension ‘thoughts’, because thoughts significantly mediate the relationship between customer confusion and customer satisfaction or customer loyalty. Analysis indicate that the R-square of thoughts was considered weak. Therefore, future studies could include more measurement items to assess thoughts and thereby increase the R-square.

5.6 Conclusion

In this study, the main focus was to examine the mediating effect of sensation in the form of gaze behavior and thoughts on the relationship between customer confusion and CX outcomes satisfaction and loyalty while participants interacted with a shopping bot. Based on our findings, customer confusion had a negative relationship with positive thoughts on the service and a positive relationship with increased gaze behavior. It was found that the effect of confusion on customer satisfaction and loyalty was significant through positive thoughts on the service, but not through increased gaze behavior. Businesses can benefit from these findings by focusing evaluations of chatbots on assessing customers thoughts on the automated service and improve their technology based on that.

Reference list

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. In *Electronic Markets* (Vol. 31, Issue 2, pp. 427–445). https://doi.org/10.1007/s12525-020-00414-7

Alreck, P. L., DiBartolo, G. R., Diriker, M., Dover, H. F., Passyn, K. A., & Settle, R. B. (2009). Time pressure, time saving and online shopping: exploring a contradiction. *Journal of Applied Business Research (JABR)*, *25*(5). https://doi.org/https://doi.org/10.19030/jabr.v25i5.1009

Ashfaq, M., Yun, J., Yu, S., & Loureiro, S. M. C. (2020). I, Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. *Telematics and Informatics*, *54*, 101473. https://doi.org/https://doi.org/10.1016/j.tele.2020.101473

Babin, B. J., & Darden, W. R. (1996). Good and bad shopping vibes: spending and patronage satisfaction. *Journal of Business Research*, *35*(3), 201–206. https://doi.org/10.1016/0148-2963(95)00125-5

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), 644–656. https://doi.org/10.1086/209376

Baier, D., Rese, A., Röglinger, M., Baier, D., Rese, A., & Röglinger, M. (2018). Conversational User Interfaces for Online Shops? A Categorization of Use Cases. *ICIS*.

Bettencourt, L. A., Lusch, R. F., & Vargo, S. L. (2014). A service lens on value creation: marketing's role in achieving strategic advantage. *California Management Review*, *57*(1), 44–66. https://doi.org/DOI: 10.1525/cmr.2014.57.1.44.

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. *Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November 22-24, 2017, Proceedings 4*, 377–392. https://doi.org/10.1007/978-3-319-70284-1_30

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons*, *63*(2), 183–193. https://doi.org/10.1016/j.bushor.2019.11.003

Chandler, J. D., & Vargo, S. L. (2011). Contextualization and value-in-context: How context frames exchange. *Marketing Theory*, *11*(1), 35–49. https://doi.org/10.1177/1470593110393713

Chen, H.-T., & Lin, Y.-T. (2018). A study of the relationships among sensory experience, emotion, and buying behavior in coffeehouse chains. *Service Business*, *12*, 551–573. https://doi.org/10.1007/s11628-017-0354-5

Chen, J.-S., Le, T.-T.-Y., & Florence, D. (2021). Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing. *International Journal of Retail & Distribution Management*, *49*(11), 1512–1531. https://doi.org/10.1108/IJRDM-08-2020-0312

Cheng, Y., & Jiang, H. (2020). How do AI-driven chatbots impact user experience? Examining gratifications, perceived privacy risk, satisfaction, loyalty, and continued use. *Journal of Broadcasting & Electronic Media*, *64*(4), 592–614. https://doi.org/10.1080/08838151.2020.1834296

Clarke, D., & Kinghorn, R. (2018). Experience is everything: Here's how to get it right. *PWC (Ed.)*, *15*.

Collier, J. E., Moore, R. S., Horky, A., & Moore, M. L. (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

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/https://doi.org/10.1007/BF02310555

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10.1207/s15326934crj1803_13

Curran, J. M., 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

De Dreu, C. K. W. (2003). Time pressure and closing of the mind in negotiation. *Organizational Behavior and Human Decision Processes*, *91*(2), 280–295. https://doi.org/https://doi.org/10.1016/S0749-5978(03)00022-0

De Keyser, A., Lemon, K. N., Klaus, P., & Keiningham, T. L. (2015). A framework for understanding and managing the customer experience. *Marketing Science Institute Working Paper Series*, *85*(1), 15–121. https://www.jstor.org/stable/26426855

De Keyser, A., Verleye, K., Lemon, K. N., Keiningham, T. L., & Klaus, P. (2020). Moving the customer experience field forward: introducing the touchpoints, context, qualities (TCQ) nomenclature. *Journal of Service Research*, *23*(4), 433–455. https://doi.org/10.1177/1094670520928390

Demming, C. L., Jahn, S., & Boztuğ, Y. (2017). Conducting Mediation Analysis in Marketing Research. *Marketing: ZFP – Journal of Research and Management*, *39*(3), 76–93. https://www.jstor.org/stable/26426855

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis*, *81*, 10–23. https://doi.org/https://doi.org/10.1016/j.csda.2014.07.008

D'mello, S., & Graesser, A. (2013). AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *2*(4), 1–39. https://doi.org/https://doi.org/10.1145/2395123.2395128

D'Mello, S., & Graesser, A. (2014). Confusion and its dynamics during device comprehension with breakdown scenarios. *Acta Psychologica*, *151*, 106–116. https://doi.org/https://doi.org/10.1016/j.actpsy.2014.06.005

Easingwood, C. J. (1986). New product development for service companies. *Journal of Product Innovation Management*, *3*(4), 264–275. https://doi.org/10.1016/0737-6782(86)90005-6

Elliott, K. M., & Hall, M. C. (2005). ASSESSING CONSUMERS'PROPENSITY TO EMBRACE SELF-SERVICE TECHNOLOGIES: ARE THERE GENDER DIFFERENCES? *Marketing Management Journal*, *15*(2).

Evelina, T. Y., Kusumawati, A., & Nimran, U. (2020). The influence of utilitarian value, hedonic value, social value, and perceived risk on customer satisfaction: survey of e-commerce customers in Indonesia. *Business: Theory and Practice*, *21*(2), 613–622.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (5th ed.). sage.

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/https://doi.org/10.1177/002224378101800010

Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding and evaluating qualitative research. *Australian & New Zealand Journal of Psychiatry*, *36*(6), 717–732.

Gelbrich, K., & Sattler, B. (2014). Anxiety, crowding, and time pressure in public self-service technology acceptance. *Journal of Services Marketing*, *28*(1), 82–94. https://doi.org/10.1108/JSM-02-2012-0051

Granka, L., Feusner, M., & Lorigo, L. (2008). Eye monitoring in online search. *Passive Eye Monitoring: Algorithms, Applications and Experiments*, 347–372. https://doi.org/10.1007/978-3-540-75412-1_16

Gujjar, J. P., & Kumar, H. P. (2021). Sentiment analysis: Textblob for decision making. *Int. J. Sci. Res. Eng. Trends*, *7*(2), 1097–1099.

Hair, F. H., Black, W. C., Babin, B. J., Anderson, R. E., Black, W. C., & Anderson, R. E. (2018). *Multivariate Data Analysis (7th Editio). Hampshire, UK Cengage Learning*.

Hoffman, D. L., & Novak, T. P. (2018). Consumer and object experience in the internet of things: An assemblage theory approach. *Journal of Consumer Research*, *44*(6), 1178–1204. https://doi.org/10.1093/jcr/ucx105

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/https://doi.org/10.1080/10705519909540118

Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology*, *22*(4), 582–594. https://doi.org/10.1016/j.jcps.2012.03.006

IBM SPSS Statistics. (2020). *IBM SPSS Statistics Versie 27: Wat is er bijgekomen?* https://www.ibm.com/blogs/think/nl-en/2020/06/24/ibm-spss-statistics-versie-27-wat-is-er-bijgekomen/

Jones, M. A., Reynolds, K. E., & Arnold, M. J. (2006). Hedonic and utilitarian shopping value: Investigating differential effects on retail outcomes. *Journal of Business Research*, *59*(9), 974–981. https://doi.org/10.1016/j.jbusres.2006.03.006

JuniperResearch. (2017). *Chatbots, a game changer for banking & healthcare, saving $8 billion annually by 2022*. Juniper Research. https://www.juniperresearch.com/press/chatbots-a-game-changer-for-banking-healthcare

Keiningham, T., Aksoy, L., Bruce, H. L., Cadet, F., Clennell, N., Hodgkinson, I. R., & Kearney, T. (2020). Customer experience driven business model innovation. *Journal of Business Research*, *116*, 431–440. https://doi.org/10.1016/j.jbusres.2019.08.003

KNAW, N. F. U. (2020). NWO, TO2-federatie, Vereniging Hogescholen, & VSNu (2018). *Netherlands Code of Conduct for Research Integrity*.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, *22*(3), 332–351. https://doi.org/10.1016/j.jcps.2011.08.003

LaSalle, D., & Britton, T. (2003). *Priceless: Turning ordinary products into extraordinary experiences*. Harvard Business Press.

Lee, H.-H., Chen, Z.-L., Yeh, S.-L., Hsiao, J. H., & Wu, A.-Y. (2021). When eyes wander around: mind-wandering as revealed by eye movement analysis with hidden Markov models. *Sensors*, *21*(22), 7569. https://doi.org/https://doi.org/10.3390/s21227569

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.

Lindstrom, M. (2005). Broad sensory branding. *Journal of Product & Brand Management*. https://doi.org/10.1108/10610420510592554

Lu, A. C. C., & Gursoy, D. (2015). A conceptual model of consumers' online tourism confusion. *International Journal of Contemporary Hospitality Management*, *27*(6), 1320–1342. https://doi.org/10.1108/IJCHM-04-2014-0171

Machleit, K. A., Eroglu, S. A., & Mantel, S. P. (2000). Perceived retail crowding and shopping satisfaction: what modifies this relationship? *Journal of Consumer Psychology*, *9*(1), 29–42. https://doi.org/10.1207/s15327663jcp0901_3

Marshall, S. P. (2002). The index of cognitive activity: Measuring cognitive workload. *Proceedings of the IEEE 7th Conference on Human Factors and Power Plants*, 7. https://doi.org/10.1109/HFPP.2002.1042860

Maule, A. J., Hockey, G. R. J., & Bdzola, L. (2000). Effects of time-pressure on decision-making under uncertainty: changes in affective state and information processing strategy. *Acta Psychologica*, *104*(3), 283–301. https://doi.org/https://doi.org/10.1016/S0001-6918(00)00033-0

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

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

Meta. (n.d.). *Meta datacenters*. Retrieved 24 June 2023, from https://datacenters.atmeta.com/europe/

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

Meuter, M. L., Ostrom, A. L., Roundtree, R. I., & Bitner, M. J. (2000). Self-service technologies: understanding customer satisfaction with technology-based service encounters. *Journal of Marketing*, *64*(3), 50–64. https://doi.org/10.1509/jmkg.64.3.50.18024

Meyer, C., & Schwager, A. (2007). Understanding customer experience. *Harvard Business Review*, *85*(2), 116.

Mindbrowser. (2020). *Chatbot Survey 2020*. https://www.mindbowser.com/chatbot-market-survey-2022/

Mitchell, V., & 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.-W., Walsh, G., & Yamin, M. (2005). Towards a conceptual model of consumer confusion. *ACR North American Advances*, *32*, 143–150.

Molla, A., & Licker, P. (2001). E-Commerce Systems Success: An Attempt to Extend and Respecify the Delone and MaClean Model of IS Success. *J. Electron. Commerce Res.*, *2*, 131–141.

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods*, *22*(1), 6–27. https://doi.org/10.1037/met0000086

Nguyen, Q. N., Sidorova, A., & Torres, R. (2022). User interactions with chatbot interfaces vs. Menu-based interfaces: An empirical study. *Computers in Human Behavior*, *128*, 107093. https://doi.org/10.1016/j.chb.2021.107093

Oliver, R. L. (1993). A conceptual model of service quality and service satisfaction: Comparative goals, different concepts. *Advances in Service Marketing and Management*, *2*, 65–85.

Oliver, R. L. (2014). *Satisfaction: A behavioral perspective on the consumer*. Routledge.
Pachman, M., Arguel, A., Lockyer, L., Kennedy, G., & Lodge, J. (2016a). Eye tracking and early detection of confusion in digital learning environments: Proof of concept. *Australasian Journal of Educational Technology*, *32*(6). https://doi.org/https://doi.org/10.14742/ajet.3060
Pachman, M., Arguel, A., Lockyer, L., Kennedy, G., & Lodge, J. (2016b). Eye tracking and early detection of confusion in digital learning environments: Proof of concept. *Australasian Journal of Educational Technology*, *32*(6).
Parasuraman, A., & Grewal, D. (2000). The Impact of Technology on the Quality-Value-Loyalty Chain: A Research Agenda. *Journal of the Academy of Marketing Science*, *28*(1), 168–174. https://doi.org/10.1177/0092070300281015
Perkhofer, L., & Lehner, O. (2019). Using gaze behavior to measure cognitive load. *Information Systems and Neuroscience: NeuroIS Retreat 2018*, 73–83. https://doi.org/https://doi.org/10.1007/978-3-030-01087-4_9
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/https://doi.org/10.1016/S0167-8116(98)00022-6
Przegalinska, A., Ciechanowski, L., Stroz, A., Gloor, P., & Mazurek, G. (2019). In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, *62*(6), 785–797. https://doi.org/10.1016/j.bushor.2019.08.005
Pupil Labs. (n.d.). *Pupil labs Core* . Retrieved 22 June 2023, from https://pupil-labs.com/products/core/
PupilLabs. (n.d.). *Homepage*. https://pupil-labs.com/products/core/
Ramasamy, A. (2019). *How businesses can begin using Chatbots the right way*. https://www.forbes.com/sites/forbestechcouncil/2019/01/18/how-businesses-can-begin-usingchatbots-the-right-way/#3d8366b12183.
Raptis, G. E., Katsini, C., Belk, M., Fidas, C., Samaras, G., & Avouris, N. (2017). Using eye gaze data and visual activities to infer human cognitive styles: method and feasibility studies. *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, 164–173. https://doi.org/https://doi.org/10.1145/3079628.3079690
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/https://doi.org/10.5898/JHRI.1.1.Riek
Rietz, T., Benke, I., & Maedche, A. (2019). *The impact of anthropomorphic and functional chatbot design features in enterprise collaboration systems on user acceptance*.
Rowley, J. (2005). Building brand webs: Customer relationship management through the Tesco Clubcard loyalty scheme. *International Journal of Retail & Distribution Management*. https://doi.org/10.1108/09590550510588361
Ryu, K., Han, H., & Jang, S. (2010). Relationships among hedonic and utilitarian values, satisfaction and behavioral intentions in the fast-casual restaurant industry. *International Journal of Contemporary Hospitality Management*, *22*(3), 416–432. https://doi.org/doi.org/10.1108/09596111011035981
Salminen, J., Jansen, B. J., An, J., Jung, S.-G., Nielsen, L., & Kwak, H. (2018). Fixation and confusion: Investigating eye-tracking participants' exposure to information in personas. *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, 110–119.

Schmitt, B. (1999). Experiential marketing. *Journal of Marketing Management*, *15*(1–3), 53–67. https://doi.org/10.1362/026725799784870496
Schweizer, M., Kotouc, A. J., & Wagner, T. (2006). Scale development for consumer confusion. *Advances in Consumer Research*, *33*(1), 184–190.
Sharma, A., Pandher, J. S., & Prakash, G. (2023). Consumer confusion and decision postponement in the online tourism domain: the moderating role of self-efficacy. *Journal of Hospitality and Tourism Insights*, *6*(2), 1092–1117.
Shaw, C., & Ivens, J. (2002). *Building great customer experiences* (Vol. 241). Springer. https://doi.org/10.1057/9780230554719
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
Sprotles, G. B., & Kendall, E. L. (1986). A methodology for profiling consumers' decision-making styles. *Journal of Consumer Affairs*, *20*(2), 267–279. https://doi.org/https://doi.org/10.1111/j.1745-6606.1986.tb00382.x
Srivastava, M., & Kaul, D. (2014). Social interaction, convenience and customer satisfaction: The mediating effect of customer experience. *Journal of Retailing and Consumer Services*, *21*(6), 1028–1037. https://doi.org/10.1016/j.jretconser.2014.04.007
Steichen, B., Fu, B., & Nguyen, T. (2020). Inferring cognitive style from eye gaze behavior during information visualization usage. *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, 348–352. https://doi.org/https://doi.org/10.1145/3340631.3394881
Textblob. (n.d.). *You can treat TextBlob objects as if they were Python strings that learned how to do Natural Language Processing.* Retrieved 29 June 2023, from https://textblob.readthedocs.io/en/dev/quickstart.html#sentiment-analysis
Trivedi, J. (2019). Examining the customer experience of using banking chatbots and its impact on brand love: The moderating role of perceived risk. *Journal of Internet Commerce*, *18*(1), 91–111. https://doi.org/10.1080/15332861.2019.1567188
Turnbull, P., Leek, S., & Ying, G. (2000). Customer Confusion: The Mobile Phone Market. *Journal of Marketing Management*, *16*, 143–163. https://doi.org/10.1362/026725700785100523
UFM. (n.d.). *Proactive Monitoring*. Retrieved 30 June 2023, from https://ufm.com.hk/proactive-monitoring/
Vincent-Wayne, M., & Vassilios, P. (1999). Marketing causes and implications of consumer confusion. *The Journal of Product & Brand Management*, *8*(4), 319–342.
Walsh, G., & Mitchell, V. (2010a). 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
Walsh, G., & Mitchell, V.-W. (2010b). The effect of consumer confusion proneness on word of mouth, trust, and customer satisfaction. *European Journal of Marketing*, *44*, 838–859. https://doi.org/10.1108/03090561011032739
Werts, C. E., Rock, D. R., Linn, R. L., & Jöreskog, K. G. (1978). A general method of estimating the reliability of a composite. *Educational and Psychological Measurement*, *38*(4), 933–938.
Whatsapp. (n.d.). *Whatsapp business*. Retrieved 23 June 2023, from https://business.whatsapp.com
Wix.com. (n.d.). *Homepage*. Retrieved 23 June 2023, from https://www.wix.com/

Zakaria, I., Rahman, B. A., Othman, A. K., Yunus, N. A. M., Dzulkipli, M. R., & Osman, M. A. F. (2014). The relationship between loyalty program, customer satisfaction and customer loyalty in retail industry: A case study. *Procedia-Social and Behavioral Sciences*, *129*, 23–30. https://doi.org/10.1016/j.sbspro.2014.03.643

Zhu, Z., Nakata, C., Sivakumar, K., & Grewal, D. (2007). Self-service technology effectiveness: the role of design features and individual traits. *Journal of the Academy of Marketing Science*, *35*, 492–506. https://doi.org/10.1007/s11747-007-0019-3

Appendix 1: 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 Tao Zhu (s1083709, Master student, Business Administration Nijmegen School of Management), email: tao.zhu@ru.nl.

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 2: Experiment conditions on WIX.com

Figure 2.1 Starting screen

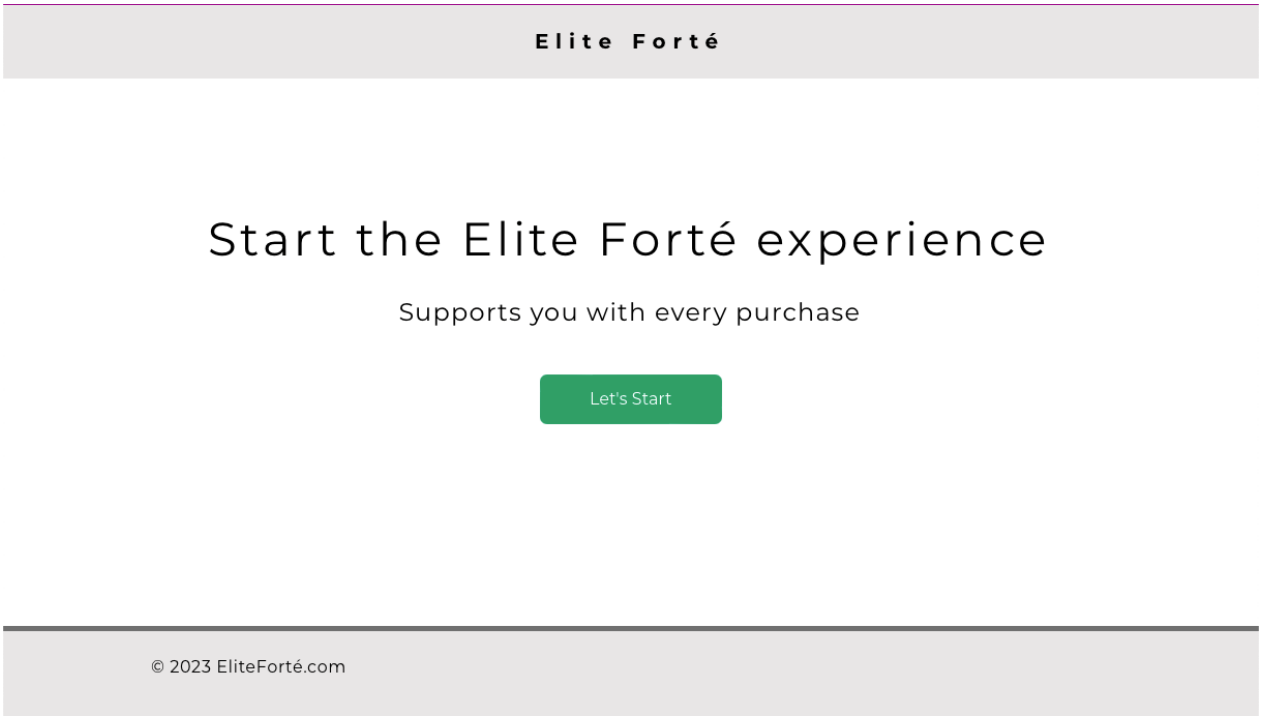


Figure 2.2 Shopping page

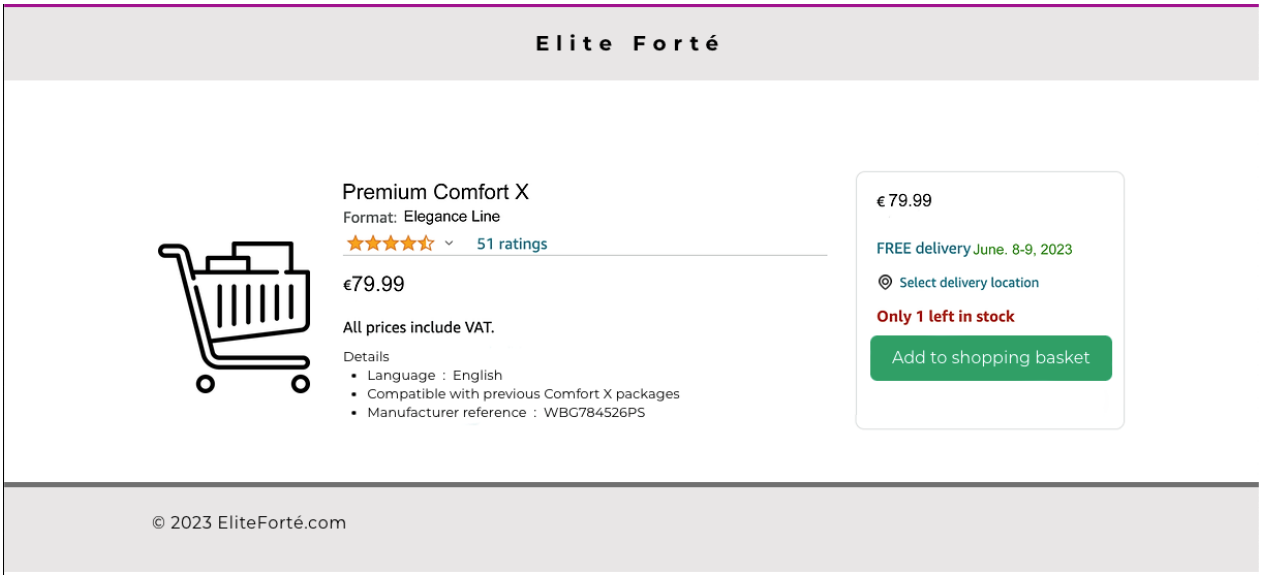


Figure 2.3 Error page condition 1 (No confusion / No time urgency)

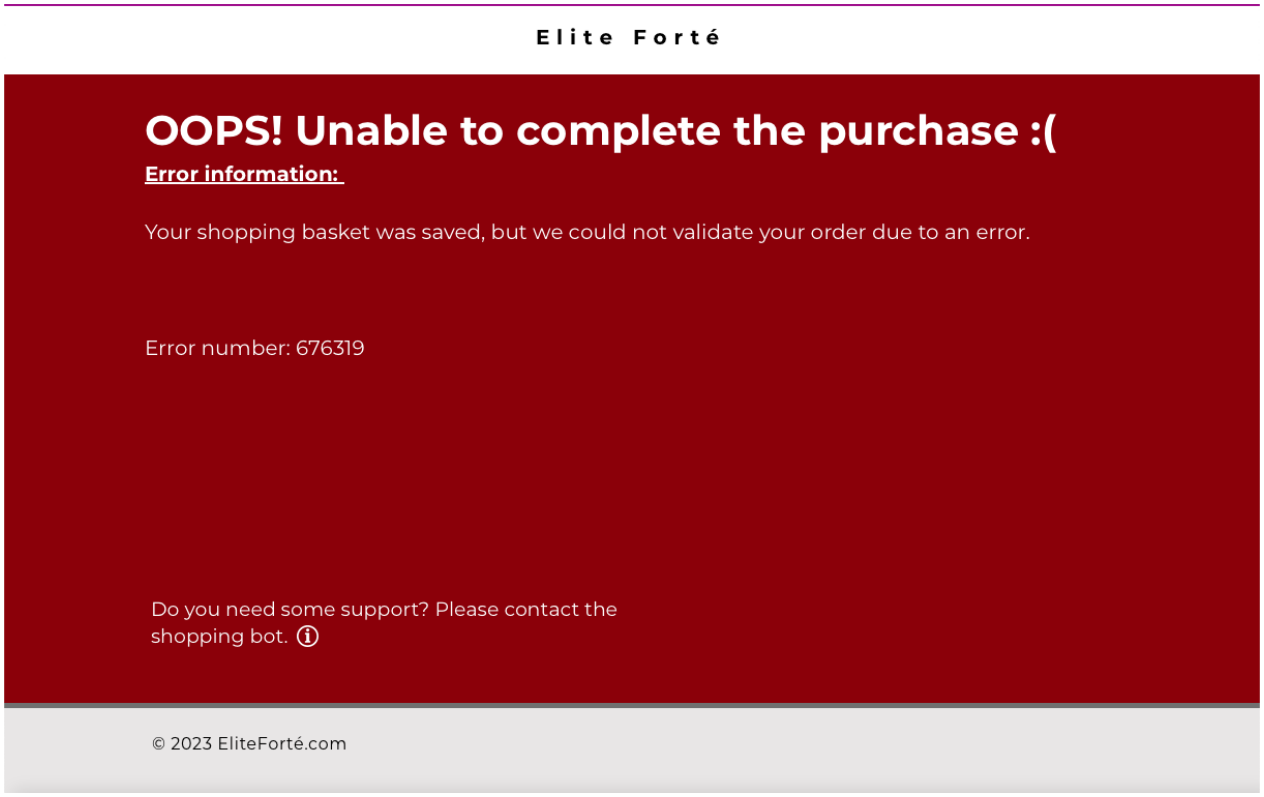


Figure 2.4 Error page condition 1 (Yes confusion / No time urgency)

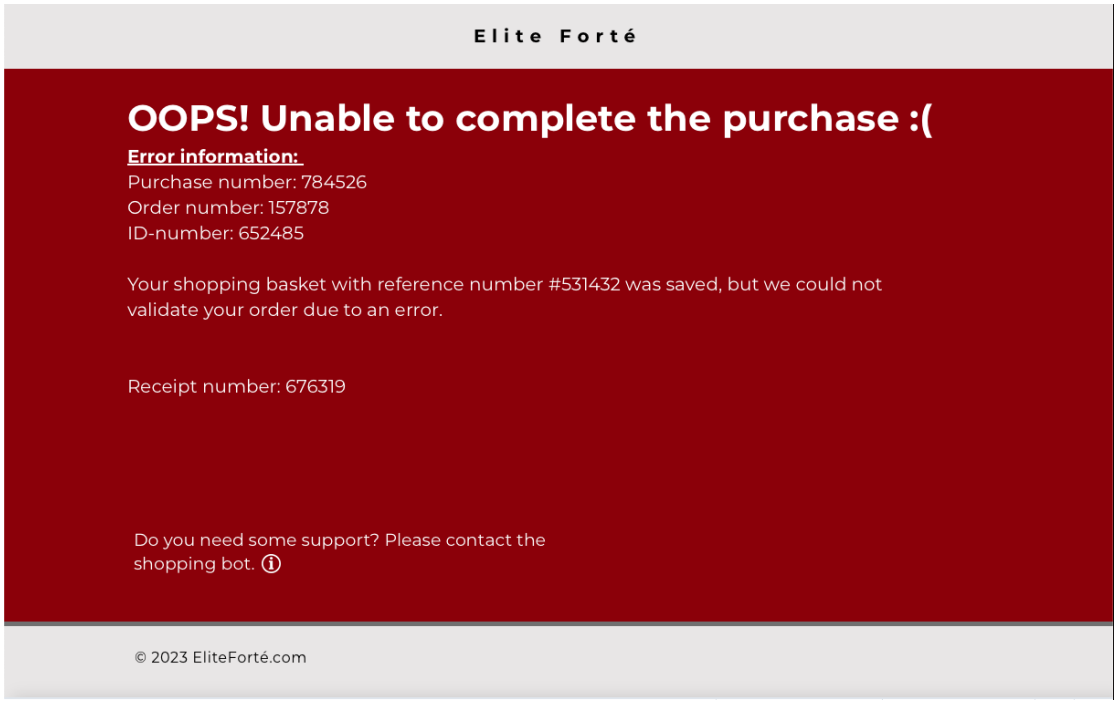


Figure 2.5 Error page condition 1 (No confusion / Yes time urgency)

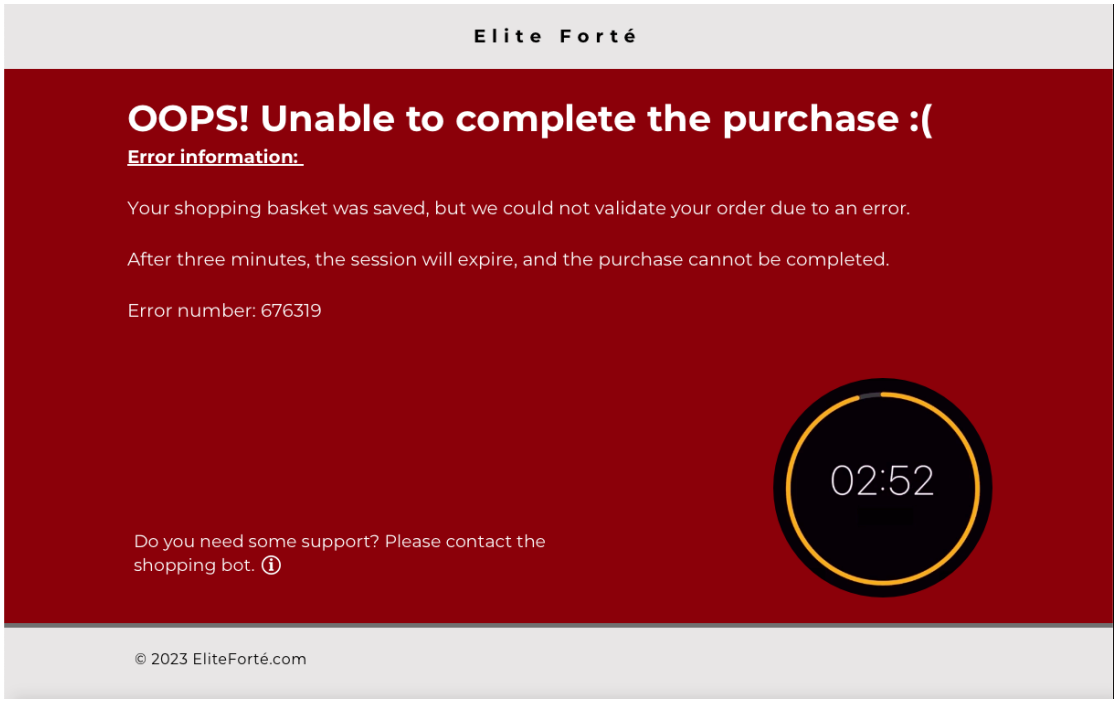


Figure 2.6 Error page condition 1 (Yes confusion / Yes time urgency)

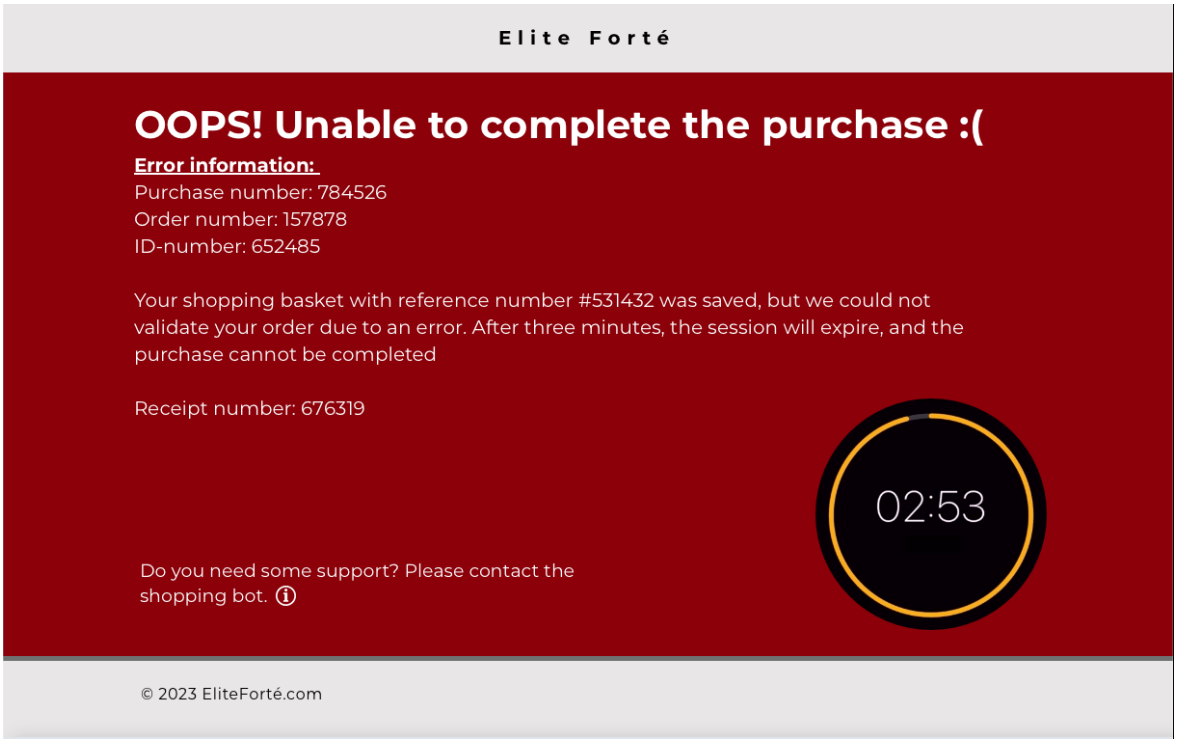
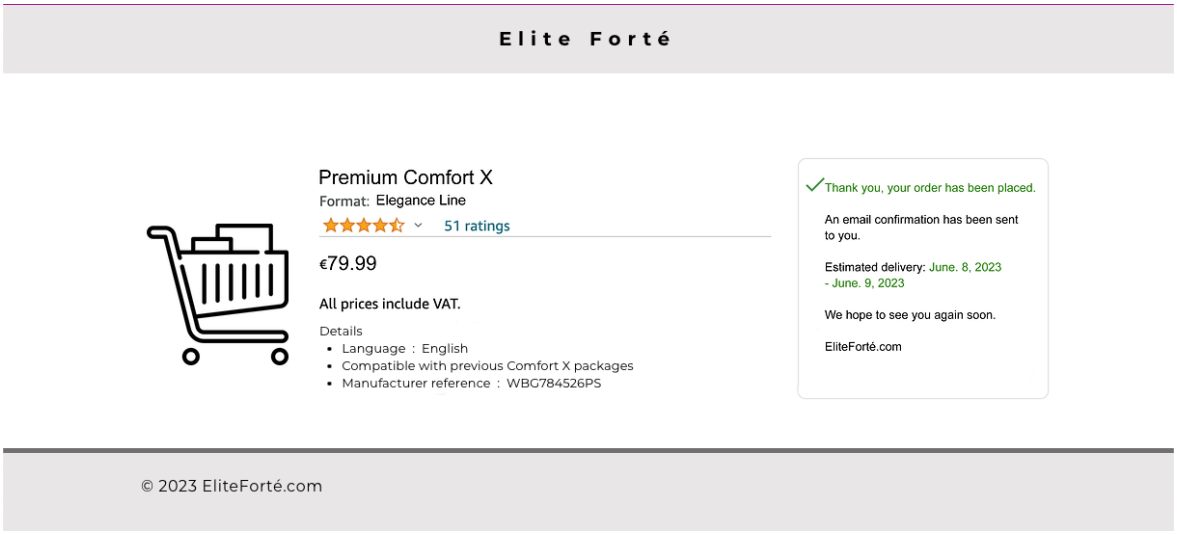


Figure 2.7 Fixed page



Appendix 3: Pre-test results

Results Pre-test for confusion

Table 3.1 Group Statistics Confusion

Group Statistics

	C_Confusion	N	Mean	Std. Deviation	Std. Error Mean
Q1_The more information I sought, the harder it seems to complete my purchase.	0	10	3.10	1.524	.482
	1	10	5.00	1.155	.365
Q2_I find the provided information clear enough to complete the purchase.	0	10	4.40	1.838	.581
	1	10	3.10	1.729	.547
Q3_The information provided to me was too ambiguous that often I feel confused.	0	10	2.30	1.567	.496
	1	10	5.70	1.160	.367
Q4_All the information I sought, makes completing my purchase less confusing.	0	10	5.50	.972	.307
	1	10	3.30	1.160	.367

Table 3.2 Group means Confusion

Group Statistics

	C_Confusion	N	Mean	Std. Deviation	Std. Error Mean
Mean_confusion	0	10	2.8750000	1.00173461	.31677630
	1	10	5.0750000	.77325071	.24452335

Table 3.3 Independent sample T-Test mean Confusion

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Mean_confusion	Equal variances assumed	.129	.724	-5.498	18	<.001	-2.2000000	.40017357	-3.0407335	-1.3592665
	Equal variances not assumed			-5.498	16.915	<.001	-2.2000000	.40017357	-3.0446153	-1.3553847

Table 3.4 Independent sample T-Test manipulation items

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Q1_The more information I sought, the harder it seems to complete my purchase.	Equal variances assumed	2.681	.119	-3.143	18	.006	-1.900	.605	-3.170	-.630
	Equal variances not assumed			-3.143	16.773	.006	-1.900	.605	-3.177	-.623
Q2_I find the provided information clear enough to complete the purchase. (Reversed)	Equal variances assumed	.180	.677	1.629	18	.121	1.300	.798	-.376	2.976
	Equal variances not assumed			1.629	17.933	.121	1.300	.798	-.377	2.977
Q3_The information provided to me was too ambiguous that often I feel confused.	Equal variances assumed	.097	.759	-5.516	18	<.001	-3.400	.616	-4.695	-2.105
	Equal variances not assumed			-5.516	16.582	<.001	-3.400	.616	-4.703	-2.097
Q4_All the information I sought, makes completing my purchase less confusing. (Reversed)	Equal variances assumed	.462	.505	4.598	18	<.001	2.200	.478	1.195	3.205
	Equal variances not assumed			4.598	17.467	<.001	2.200	.478	1.193	3.207

Table 3.5 One-way ANOVA manipulation items confusion

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Q1_The more information I sought, the harder it seems to complete my purchase.	Between Groups	18.050	1	18.050	9.875	.006
	Within Groups	32.900	18	1.828		
	Total	50.950	19			
Q2_I find the provided information clear enough to complete the purchase.	Between Groups	8.450	1	8.450	2.654	.121
	Within Groups	57.300	18	3.183		
	Total	65.750	19			
Q3_The information provided to me was too ambiguous that often I feel confused.	Between Groups	57.800	1	57.800	30.421	<.001
	Within Groups	34.200	18	1.900		
	Total	92.000	19			
Q4_All the information I sought, makes completing my purchase less confusing.	Between Groups	24.200	1	24.200	21.146	<.001
	Within Groups	20.600	18	1.144		
	Total	44.800	19			

Results Pre-test for Time urgency

Table 3.6 Group Statistics **Time urgency**

Group Statistics

	Condition number	N	Mean	Std. Deviation	Std. Error Mean
Q1_I felt like I had enough time to complete the purchase.	0	10	2.40	1.265	.400
	1	10	3.90	2.183	.690
Q2_I felt I was under time pressure while making the purchase.	0	10	3.00	1.563	.494
	1	10	4.90	2.183	.690
Q3_The presence of a timer made me feel rushed during the purchase.	0	10	3.50	1.958	.619
	1	10	5.80	2.300	.727

Table 3.7 Group means **Time urgency**

Group Statistics

	Condition number	N	Mean	Std. Deviation	Std. Error Mean
Mean_TimePressure	0	10	2.9666667	1.05934991	.33499585
	1	10	4.8666667	1.96387121	.62103060

Table 3.8 Independent sample T-Test mean **Time urgency**

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Mean_TimePressure	Equal variances assumed	2.207	.155	-2.693	18	.015	-1.9000000	.70562117	-3.3824551	-.41754493
	Equal variances not assumed			-2.693	13.829	.018	-1.9000000	.70562117	-3.4151676	-.38483237

Table 3.9 Independent sample T-Test manipulation items **Time urgency**

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Q1_I felt like I had enough time to complete the purchase.	Equal variances assumed	6.636	.019	-1.880	18	.076	-1.500	.798	-3.176	.176
	Equal variances not assumed			-1.880	14.430	.080	-1.500	.798	-3.207	.207
Q2_I felt I was under time pressure while making the purchase.	Equal variances assumed	.692	.416	-2.237	18	.038	-1.900	.849	-3.684	-.116
	Equal variances not assumed			-2.237	16.309	.040	-1.900	.849	-3.697	-.103
Q3_The presence of a timer made me feel rushed during the purchase.	Equal variances assumed	.048	.829	-2.408	18	.027	-2.300	.955	-4.307	-.293
	Equal variances not assumed			-2.408	17.553	.027	-2.300	.955	-4.310	-.290

Table 3.10 One-way ANOVA manipulation items **Time urgency**

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Q1_I felt like I had enough time to complete the purchase. (Reversed)	Between Groups	11.250	1	11.250	3.534	.076
	Within Groups	57.300	18	3.183		
	Total	68.550	19			
Q2_I felt I was under time pressure while making the purchase.	Between Groups	18.050	1	18.050	5.006	.038
	Within Groups	64.900	18	3.606		
	Total	82.950	19			
Q3_The presence of a timer made me feel rushed during the purchase.	Between Groups	26.450	1	26.450	5.799	.027
	Within Groups	82.100	18	4.561		
	Total	108.550	19			

Appendix 4: Operationalization table

Table 4.1 Opeationalisation table

Concept	Measurement description	Source
Sensations		
Gaze behavior	Total number of fixations Average duration of fixations	Meißner & Oll (2019), Pachman et al. (2016), Salminen et al. (2018)
Thoughts		
	- Please take me through what you were thinking as you were attempting to help me resolve the purchase error we encountered. - What actions crossed your mind in the process of resolving this error? - Would you please elaborate by telling me what was valuable in achieving the outcome of the shopping experience and how?	-
Utilitarian value	- To what extent do you agree that we have accomplished what we wanted from this purchase? - 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?	Babin et al. (1994)
Customer experience outcomes		
Satisfaction	- So as an outcome, to what extent would you say that you are satisfied with the shopping experience? (Strongly Dissatisfied – Strongly Satisfied)	-
Loyalty	- 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)

Concept	Measurement description	Source
Control variables		
Age	- Can you please provide me with your age?	-
Gender	- What is the gender you identify with the most? (1) Male, (2) Female, (3) Other, (4) Prefer not to say	-
Tolerance for ambiguity	- To what extent do you agree that you prefer situations in which there is some ambiguity? - 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?	McLain (2009)
Confusion Avoidance	- To what extent do you agree that providing clear information is the basis for avoiding confusion? - 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?	Schweizer et al. (2006)
Manipulation Check		
Confusion	- 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)
Time pressure	- 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)

Concept	Measurement description	Source
	- To what extent do you agree that the presence of a timer made you feel rushed during the purchase?	
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?	-

Appendix 5: Results measurement and structural model

Table 5.1 Univariate statistics

Univariate Statistics

	N	Mean	Std. Deviation	Missing		No. of Extremes[a]	
				Count	Percent	Low	High
TfAQ1	197	3.44	1.670	0	.0	0	0
TfAQ2	197	4.73	1.391	0	.0	6	0
TfAQ3	197	4.42	1.328	0	.0	0	0
Age	196	25.17	8.186	1	.5	0	18
Totalfixationsrelevantareas	197	2143.70	1077.982	0	.0	0	1
Averagedurationfixationsrelevantareas	197	168.335787	23.3768402	0	.0	0	4
Totalfixationsnonrelevantareas	197	418.82	419.399	0	.0	0	13
Averagedurationfixationsnonrelevantareas	197	167.079	25.2329	0	.0	0	1
ThoughtsOQ1	197	4.31767114	1.07633826	0	.0	2	0
ThoughtsOQ2	197	4.69450948	.886140680	0	.0	11	11
ThoughtsOQ3	197	3.12635067	.986745160	0	.0	0	2
ThoughtsCQ1	197	5.74	1.385	0	.0	7	0
ThoughtsCQ2	196	5.70	1.628	1	.5	16	0
ThoughtsCQ3	197	5.36	1.686	0	.0	19	0
ThoughtsCQ4	197	5.05	1.576	0	.0	10	0
Satisfaction	196	5.04	1.424	1	.5	4	0
Loyalty	197	4.54	1.768	0	.0	0	0
NPS	197	4.47	1.704	0	.0	0	0
Alternativeexplainabilityampcontrols Q1	196	5.76	1.215	1	.5	7	0
Alternativeexplainabilityampcontrols Q2	197	2.49	1.649	0	.0	0	0
Alternativeexplainabilityampcontrols Q3	196	3.98	1.799	1	.5	0	0
CAQ1	197	6.64	.786	0	.0	5	0
CAQ2	197	5.86	1.535	0	.0	11	0
CAQ3	197	3.9340	1.55534	0	.0	0	0
Gender	197			0	.0		

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

Table 5.2 EM Means

EM Means[a]

TfAQ1	TfAQ2	TfAQ3	Age	Totalfixationsrelevantareas	Averagedurationfixationsrelevantareas	Totalfixationsnonrelevantareas	Averagedurationfixationsnonrelevantareas	ThoughtsOQ1	ThoughtsOQ2	ThoughtsOQ3	ThoughtsCQ1	ThoughtsCQ2	ThoughtsCQ3	ThoughtsCQ4	Satisfaction	Loyalty	NPS	Alternativeexplainabilityampcontrols Q1	Alternativeexplainabilityampcontrols Q2	Alternativeexplainabilityampcontrols Q3	CAQ1	CAQ2	CAQ3
3.44	4.73	4.42	25.20	2143.70	168.335787	418.82	167.079	4.31767114	4.69450948	3.12635067	5.74	5.71	5.36	5.05	5.05	4.54	4.47	5.76	2.49	3.99	6.64	5.86	3.9340

a. Little's MCAR test: Chi-Square = 139,194, DF = 115, Sig. = ,062

Manipulation check

Table 5.3 Group statistics Confusion

Group Statistics

	Confusion	N	Mean	Std. Deviation	Std. Error Mean
MC Confusion Q1	0	99	3.59	1.756	.176
	1	98	5.22	1.297	.131
MC Confusion Q3	0	99	3.52	1.961	.197
	1	98	5.73	1.328	.134
MCConfusionQ2R	0	98	2.8571	1.89954	.19188
	1	98	4.9286	1.73651	.17541
MCConfusionQ4R	0	99	3.1313	1.56261	.15705
	1	98	4.1633	1.49703	.15122

Table 5.4 Independent sample T-Test Confusion

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
MC Confusion Q1	Equal variances assumed	18.636	<.001	-7.446	195	<.001	-1.639	.220	-2.073	-1.205
	Equal variances not assumed			-7.457	180.414	<.001	-1.639	.220	-2.072	-1.205
MC Confusion Q3	Equal variances assumed	43.250	<.001	-9.293	195	<.001	-2.220	.239	-2.691	-1.749
	Equal variances not assumed			-9.311	172.471	<.001	-2.220	.238	-2.690	-1.749
MCConfusionQ2R	Equal variances assumed	1.181	.278	-7.968	194	<.001	-2.07143	.25998	-2.58418	-1.55868
	Equal variances not assumed			-7.968	192.459	<.001	-2.07143	.25998	-2.58420	-1.55866
MCConfusionQ4R	Equal variances assumed	.970	.326	-4.732	195	<.001	-1.03195	.21807	-1.46203	-.60188
	Equal variances not assumed			-4.733	194.792	<.001	-1.03195	.21802	-1.46194	-.60197

Table 5.5 One-way Anova Confusion

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
MC Confusion Q1	Between Groups	132.238	1	132.238	55.445	<.001
	Within Groups	465.081	195	2.385		
	Total	597.320	196			
MCConfusionQ2R	Between Groups	210.250	1	210.250	63.484	<.001
	Within Groups	642.500	194	3.312		
	Total	852.750	195			
MC Confusion Q3	Between Groups	242.617	1	242.617	86.360	<.001
	Within Groups	547.829	195	2.809		
	Total	790.447	196			
MCConfusionQ4R	Between Groups	52.446	1	52.446	22.394	<.001
	Within Groups	456.681	195	2.342		
	Total	509.127	196			

Table 5.6 Group statistics Time Urgency

Group Statistics

	Urgency	N	Mean	Std. Deviation	Std. Error Mean
MCUrgencyQ1R	0	99	2.5152	1.45941	.14668
	1	97	2.6495	1.54138	.15650
MC Urgency Q2	0	100	3.51	1.878	.188
	1	97	4.19	2.012	.204
MC Urgency Q3	0	100	3.74	1.862	.186
	1	97	5.09	1.838	.187

Table 5.7 Independent Sample T-Test Time Urgency

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
MCUrgencyQ1R	Equal variances assumed	.288	.592	-.627	194	.532	-.13433	.21437	-.55713	.28847
	Equal variances not assumed			-.626	192.913	.532	-.13433	.21449	-.55739	.28872
MC Urgency Q2	Equal variances assumed	.536	.465	-2.437	195	.016	-.676	.277	-1.222	-.129
	Equal variances not assumed			-2.435	193.083	.016	-.676	.277	-1.223	-.128
MC Urgency Q3	Equal variances assumed	.510	.476	-5.131	195	<.001	-1.353	.264	-1.873	-.833
	Equal variances not assumed			-5.132	194.941	<.001	-1.353	.264	-1.873	-.833

Table 5.8 One-way ANOVA Time Urgency

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
MCUrgencyQ1R	Between Groups	,884	1	,884	,393	,532
	Within Groups	436,810	194	2,252		
	Total	437,694	195			
MC Urgency Q2	Between Groups	22,472	1	22,472	5,941	,016
	Within Groups	737,650	195	3,783		
	Total	760,122	196			
MC Urgency Q3	Between Groups	90,108	1	90,108	26,327	<,001
	Within Groups	667,405	195	3,423		
	Total	757,513	196			

Evaluation of the measurement model

Table 5.9 Model fit

	Saturated model	Estimated model
SRMR	0.039	0.127
d_ULS	0.317	3.398
d_G	0.125	n/a
Chi-square	138.777	n/a
NFI	0.875	n/a

Construct reliability and validity

Table 5.10 Initial outloading

	Outer loadings	
Age <- Age		1.000
AlternativeexplainabilityampcontrolsQ1 <- ALT Explain Q1		1.000
AlternativeexplainabilityampcontrolsQ2 <- Alt explain Q2		1.000
AlternativeexplainabilityampcontrolsQ3 <- Alt explain Q3		1.000
Averagedurationfixationsnonrelevantareas <- Fix. duration NON-RA		1.000
Averagedurationfixationsrelevantareas <- Fix. Duration RA		1.000
CAQ1 <- CAQ	-0.467	
CAQ2 <- CAQ	0.557	
CAQ3 <- CAQ	0.643	
Confusion <- Confusion		1.000
Gender <- Gender		1.000
Loyalty <- Loyalty	0.931	
NPS <- Loyalty	0.724	
Satisfaction <- Satisfaction		1.000
TfAQ1 <- TfAQ	0.480	
TfAQ2 <- TfAQ	-0.528	
TfAQ3 <- TfAQ	0.330	
ThoughtsCQ1 <- ThoughtsCQ	0.721	
ThoughtsCQ2 <- ThoughtsCQ	0.771	
ThoughtsCQ3 <- ThoughtsCQ	0.751	
ThoughtsCQ4 <- ThoughtsCQ	0.887	
ThoughtsOQ1 <- ThoughtsOQ	0.858	
ThoughtsOQ2 <- ThoughtsOQ	0.440	
ThoughtsOQ3 <- ThoughtsOQ	0.393	
Totalfixationsnonrelevantareas <- Fixation NON-RA		1.000

	Outer loadings
Totalfixationsrelevantareas <- Fixation RA	1.000
Urgency <- Urgency	1.000
Urgency x Confusion -> Urgency x Confusion	1.000

Table 5.11 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)		Composite reliability (rho_c)	Average variance extracted (AVE)
CAQ	0.080	-0.096		0.207	0.314
Loyalty	0.593	0.754		0.818	0.695
TfAQ	0.612		-1.121	0.032	0.206
ThoughtsCQ	0.791	0.816		0.865	0.617
ThoughtsOQ	0.173	0.177		0.599	0.362

Decision in this step? Take a first look at the constructs CAQ, ThoughtsOQ and TfaQ, because the threshold for Cronbach's alpha >.700 and the Composite reliability >.700. Only question 2 loads on the construct; the other questions are very low; it is possible that questions 1 and 3 measure something different from question 2 given that they are both negative and low; therefore, you might want to remove question 2.

Table 5.12 Outer loadings

	Outer loadings
TfAQ1 <- TfAQ	0.877
TfAQ3 <- TfAQ	0.721

Table 5.13 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
TfAQ	0.460	0.498	0.782	0.644

TfAQ1 & TfAQ3 loads much beter after deleting. However the threshold for *Cronbach's alpha >.700 and the Composite reliability >.700. are not met. Therefore, The items with the lowest loading was removed.*

Table 5.14 Outer loadings

	Outer loadings
CAQ1 <- CAQ	-0.425
CAQ2 <- CAQ	0.632
CAQ3 <- CAQ	0.602

Table 5.15 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
CAQ	0.080	-0.096	0.242	0.314
ThoughtsCQ	0.791	0.817	0.865	0.616
ThoughtsOQ	0.173	0.179	0.593	0.360

Decision in this step? Take a first look at the construct CAQ, because the threshold for every criteria is not met. Item one of CAQ load negatively and does not load very well well on the construct. Therefore, CAQ1 is deleted.

Table 5.16 Outer loadings

	Outer loadings
CAQ2 <- CAQ	0.730
CAQ3 <- CAQ	0.679

Table 5.17 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
CAQ	-0.012	-0.012	0.664	0.497

Decision in this step? The construct CAQ does still violate all the thresholds. Therefore, item 3 of CAQ is deleted.

Table 5.18 Outer loadings

	Outer loadings
ThoughtsCQ1 <- ThoughtsCQ	0.724
ThoughtsCQ2 <- ThoughtsCQ	0.769
ThoughtsCQ3 <- ThoughtsCQ	0.749
ThoughtsCQ4 <- ThoughtsCQ	0.888
ThoughtsOQ1 <- ThoughtsOQ	0.858
ThoughtsOQ2 <- ThoughtsOQ	0.427
ThoughtsOQ3 <- ThoughtsOQ	0.408

Table 5.19 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
ThoughtsCQ	0.791	0.817	0.865	0.616
ThoughtsOQ	0.173	0.175	0.599	0.361

Decision in this step? The construct Thouhts open questions and thoughts closed questions have multiple items. However, the outerloadings show that there are no really high loadings. In addition, all the thresholds of the construct reliability and validity are not met. Therefore, the construct thoughts open questions is examined. Based on the loadings, ThoughtsOQ3 was removed.

Table 5.20 Outer loadings

	Outer loadings
ThoughtsOQ1 <- ThoughtsOQ	0.927
ThoughtsOQ2 <- ThoughtsOQ	0.426

Table 5.21 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
ThoughtsCQ	0.791	0.817	0.865	0.616
ThoughtsOQ	0.103	0.148	0.656	0.520

Decision in this step? Thoughts open questions still violates almost all threshols, excepts the threshold of the AVE > .500. Therefore, ThoughtsOQ2 will be removed.

Table 5.22 Outer loadings

	Outer loadings
ThoughtsCQ1 <- ThoughtsCQ	0.724
ThoughtsCQ2 <- ThoughtsCQ	0.769
ThoughtsCQ3 <- ThoughtsCQ	0.749
ThoughtsCQ4 <- ThoughtsCQ	0.888
ThoughtsOQ1 <- ThoughtsOQ	1.000

Table 5.23 Construct reliability and validity.

	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
ThoughtsCQ	0.791	0.817	0.865	0.616

Decision in this step? Thoughts closed questions items are loading well and also all thresholds of the construct reliability and validity is met. In short, Thoughts CQ construct is acceptable.

Table 5.24 Fornell-Lacker criterion

	ALT Explain Q1	Age	Alt explain Q2	Alt explain Q3	CAQ	Confusion	Fix. Duration RA	Fix. duration NON-RA	Fixation NON-RA
ALT Explain Q1	1.000								
Age	-0.131	1.000							
Alt explain Q2	-0.081	0.065	1.000						
Alt explain Q3	0.068	0.201	0.109	1.000					
CAQ	-0.016	0.117	-0.051	0.068	1.000				
Confusion	-0.029	0.052	-0.116	0.292	-0.054	1.000			
Fix. Duration RA	0.093	-0.200	0.095	0.025	0.009	-0.085	1.000		
Fix. duration NON-RA	0.096	-0.123	0.114	-0.055	-0.038	0.040	0.532	1.000	
Fixation NON-RA	-0.112	0.076	0.119	0.138	0.108	0.010	0.026	0.152	1.000
Fixation RA	-0.009	0.054	-0.002	0.157	0.007	0.190	0.237	-0.008	-0.232
Gender	0.013	0.072	0.025	-0.088	-0.065	-0.004	0.101	0.156	-0.128
Loyalty	0.080	-0.260	0.022	-0.375	-0.075	-0.223	0.111	0.072	-0.013
Satisfaction	0.092	-0.263	0.063	-0.414	-0.103	-0.315	0.087	0.030	-0.015
TfAQ	0.058	-0.015	0.010	-0.073	-0.262	-0.081	0.032	0.065	-0.137
ThoughtsCQ	0.019	-0.224	0.008	-0.457	-0.032	-0.391	0.045	0.058	-0.003
ThoughtsOQ	0.034	-0.053	0.149	-0.163	-0.127	-0.163	0.044	0.021	0.017
Urgency	0.042	-0.114	0.116	-0.060	-0.001	-0.005	0.046	0.062	0.011

Fixation RA	Gender	Loyalty	Satisfaction	TfAQ	ThoughtsCQ	ThoughtsOQ	Urgency
1.000							
0.043	1.000						
-0.092	0.082	1.000					
-0.118	0.039	0.779	1.000				
0.032	0.045	0.056	0.061	1.000			
-0.138	0.039	0.744	0.751	0.068	0.785		
-0.026	0.056	0.314	0.422	0.038	0.354	1.000	
-0.106	-0.013	0.114	0.122	0.068	0.126	-0.020	1.000

Table 5.25 Discriminant validty (HTMT)

	ALT Explain Q1	Age	Alt explain Q2	Alt explain Q3	CAQ	Confusion	Fix. Duration RA	Fix. duration NON-RA	Fixation NON-RA	Fixation RA
ALT Explain Q1										
Age	0.131									
Alt explain Q2	0.081	0.065								
Alt explain Q3	0.068	0.201	0.109							
CAQ	0.016	0.117	0.051	0.068						
Confusion	0.029	0.052	0.116	0.292	0.054					
Fix. Duration RA	0.093	0.200	0.095	0.025	0.009	0.085				
Fix. duration NON-RA	0.096	0.123	0.114	0.055	0.038	0.040	0.532			
Fixation NON-RA	0.112	0.076	0.119	0.138	0.108	0.010	0.026	0.152		
Fixation RA	0.009	0.054	0.002	0.157	0.007	0.190	0.237	0.008	0.232	
Gender	0.013	0.072	0.025	0.088	0.065	0.004	0.101	0.156	0.128	0.043
Loyalty	0.066	0.324	0.025	0.476	0.098	0.274	0.094	0.070	0.026	0.152
Satisfaction	0.092	0.263	0.063	0.414	0.103	0.315	0.087	0.030	0.015	0.118
TfAQ	0.058	0.015	0.010	0.073	0.262	0.081	0.032	0.065	0.137	0.032
ThoughtsCQ	0.052	0.251	0.011	0.513	0.052	0.451	0.063	0.063	0.110	0.157
ThoughtsOQ	0.034	0.053	0.149	0.163	0.127	0.163	0.044	0.021	0.017	0.026
Urgency	0.042	0.114	0.116	0.060	0.001	0.005	0.046	0.062	0.011	0.106
Urgency x Confusion	0.063	0.048	0.019	0.111	0.022	0.570	0.110	0.030	0.010	0.064

Gender	Loyalty	Satisfaction	TfAQ	ThoughtsCQ	ThoughtsOQ	Urgency	Urgency x Confusion
0.109							
0.039	0.854						
0.045	0.146	0.061					
0.078	0.909	0.829	0.116				
0.056	0.353	0.422	0.038	0.392			
0.013	0.101	0.122	0.068	0.142	0.020		
0.000	0.146	0.099	0.006	0.155	0.067	0.576	

- Values for the HTMT test should be less than 0.85. The following do not apply:
- Loyalty and satisfaction
Thoughts and loyalty
- are both slightly above the 0.85 threshold, indicating that the latent variables are largely comparable.

Decision? The lowest Loyalty score of 0.717 may therefore be better removed because it appears that the NPS is too similar to satisfaction and thoughts. After removing NPS, the HTMT is as follows:

Table 5.26 Discriminant validty (HTMT) after deletion of NPS

	ALT Explain Q1	Age	Alt explain Q2	Alt explain Q3	CAQ	Confusion	Fix. Duration RA	Fix. duration NON-RA
ALT Explain Q1								
Age	0.131							
Alt explain Q2	0.081	0.065						
Alt explain Q3	0.068	0.201	0.109					
CAQ	0.016	0.117	0.051	0.068				
Confusion	0.029	0.052	0.116	0.292	0.054			
Fix. Duration RA	0.093	0.200	0.095	0.025	0.009	0.085		
Fix. duration NON-RA	0.096	0.123	0.114	0.055	0.038	0.040	0.532	
Fixation NON-RA	0.112	0.076	0.119	0.138	0.108	0.010	0.026	0.152
Fixation RA	0.009	0.054	0.002	0.157	0.007	0.190	0.237	0.008
Gender	0.013	0.072	0.025	0.088	0.065	0.004	0.101	0.156
Loyalty	0.080	0.260	0.022	0.375	0.075	0.223	0.111	0.072
Satisfaction	0.092	0.263	0.063	0.414	0.103	0.315	0.087	0.030
TfAQ	0.058	0.015	0.010	0.073	0.262	0.081	0.032	0.065
ThoughtsCQ	0.052	0.251	0.011	0.513	0.052	0.451	0.063	0.063
ThoughtsOQ	0.034	0.053	0.149	0.163	0.127	0.163	0.044	0.021
Urgency	0.042	0.114	0.116	0.060	0.001	0.005	0.046	0.062
Urgency x Confusion	0.063	0.048	0.019	0.111	0.022	0.570	0.110	0.030

Fixation NON-RA	Fixation RA	Gender	Loyalty	Satisfaction	TfAQ	ThoughtsCQ	ThoughtsOQ	Urgency	Urgency x Confusion
0.232									
0.128	0.043								
0.013	0.092	0.082							
0.015	0.118	0.039	0.779						
0.137	0.032	0.045	0.056	0.061					
0.110	0.157	0.078	0.819	0.829	0.116				
0.017	0.026	0.056	0.314	0.422	0.038	0.392			
0.011	0.106	0.013	0.114	0.122	0.068	0.142	0.020		
0.010	0.064	0.000	0.059	0.099	0.006	0.155	0.067	0.576	

Evaluation of the structural model

Table 5.27 Collinearity (VIF values)

	Fix. Duration RA	Fix. duration NON-RA	Fixation NON-RA	Fixation RA	Gender	Loyalty	Satisfaction	TfAQ	ThoughtsCQ	ThoughtsOQ	Urgency	Urgency x Confusion
ALT Explain Q1						1.076	1.076					
Age						1.189	1.189					
Alt explain Q2						1.109	1.109					
Alt explain Q3						1.438	1.438					
CAQ						1.135	1.135					
Confusion	1.970	1.970	1.000	1.970		1.341	1.341		1.082	1.000		
Fix. Duration RA						1.657	1.657		1.576			
Fix. duration NON-RA						1.576	1.576		1.484			
Fixation NON-RA						1.218	1.218		1.086	1.000		
Fixation RA						1.256	1.256		1.219			
Gender						1.078	1.078					
Loyalty												
Satisfaction												
TfAQ						1.111	1.111					
ThoughtsCQ						1.581	1.581					
ThoughtsOQ						1.199	1.199					
Urgency	1.990	1.990		1.990								

	Fix. Duration RA	Fix. duration NON-RA	Fixation NON-RA	Fixation RA	Gender	Loyalty	Satisfaction	TfAQ	Thoughts CQ	Thoughts OQ	Urgency	Urgency x Confusion
Urgency x Confusion	2.949	2.949		2.949								

Table 5.28 R-Square (overview)

	R-square	R-square adjusted
Fix. Duration RA	0.032	0.017
Fix. duration NON-RA	0.029	0.014
Fixation NON-RA	0.000	-0.005
Fixation RA	0.048	0.033
Loyalty	0.586	0.554
Satisfaction	0.622	0.593
ThoughtsCQ	0.164	0.142
ThoughtsOQ	0.027	0.017

Table 5.29 Effect size (f2)

	ALT Explain Q1	Age	Alt explain Q2	Alt explain Q3	CAQ	Confusion	Fix. Duration RA	Fix. duration NON-RA	Fixation NON-RA	Fixation RA	Gender
ALT Explain Q1											
Age											
Alt explain Q2											
Alt explain Q3											
CAQ											
Confusion							0.002	0.019	0.000	0.014	
Fix. Duration RA											
Fix. duration NON-RA											
Fixation NON-RA											
Fixation RA											
Gender											
Loyalty											
Satisfaction											
TfAQ											
ThoughtsCQ											
ThoughtsOQ											
Urgency							0.020	0.024		0.010	
Urgency x Confusion							0.024	0.024		0.001	

Loyalty	Satisfaction	TfAQ	ThoughtsCQ	ThoughtsOQ	Urgency	Urgency x Confusion
0.009	0.015					
0.011	0.013					
0.002	0.007					
0.007	0.019					
0.002	0.006					
0.020	0.000		0.162	0.027		
0.013	0.010		0.000			
0.005	0.012		0.006			
0.001	0.001		0.001	0.000		
0.001	0.001		0.004			
0.006	0.000					
0.000	0.000					
0.784	0.665					
0.004	0.056					

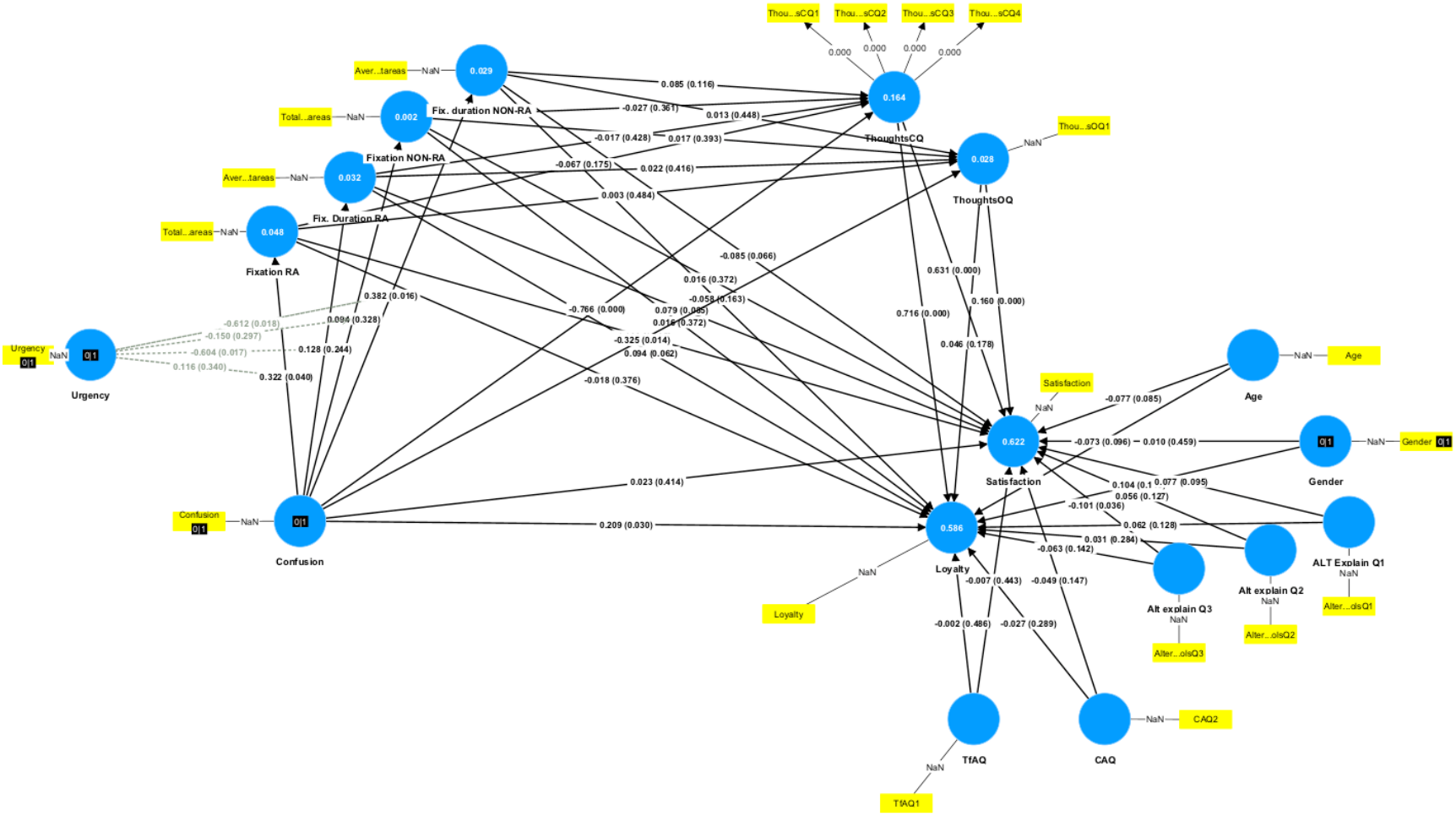


Figure 5.1 Structural model SmartPLS

Table 5.30 Total effects SmartPLS

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values
ALT Explain Q1 -> Loyalty	0.062	0.061	0.054	1.137	0.128
ALT Explain Q1 -> Satisfaction	0.077	0.075	0.059	1.312	0.095
Age -> Loyalty	-0.073	-0.074	0.056	1.306	0.096
Age -> Satisfaction	-0.077	-0.077	0.056	1.375	0.085
Alt explain Q2 -> Loyalty	0.031	0.028	0.054	0.572	0.284
Alt explain Q2 -> Satisfaction	0.056	0.054	0.049	1.139	0.127
Alt explain Q3 -> Loyalty	-0.063	-0.061	0.059	1.070	0.142
Alt explain Q3 -> Satisfaction	-0.101	-0.099	0.056	1.803	0.036
CAQ -> Loyalty	-0.027	-0.026	0.049	0.556	0.289
CAQ -> Satisfaction	-0.049	-0.046	0.046	1.052	0.147
Confusion -> Fix. Duration RA	0.128	0.131	0.185	0.692	0.244
Confusion -> Fix. duration NON-RA	0.382	0.386	0.178	2.149	0.016
Confusion -> Fixation NON-RA	0.094	0.087	0.211	0.446	0.328
Confusion -> Fixation RA	0.322	0.324	0.184	1.749	0.040
Confusion -> Loyalty	-0.364	-0.361	0.144	2.529	0.006
Confusion -> Satisfaction	-0.535	-0.530	0.127	4.216	0.000
Confusion -> ThoughtsCQ	-0.760	-0.757	0.121	6.301	0.000
Confusion -> ThoughtsOQ	-0.314	-0.326	0.145	2.164	0.015
Fix. Duration RA -> Loyalty	0.083	0.079	0.091	0.913	0.181
Fix. Duration RA -> Satisfaction	0.072	0.068	0.078	0.920	0.179
Fix. Duration RA -> ThoughtsCQ	-0.017	-0.021	0.091	0.182	0.428
Fix. Duration RA -> ThoughtsOQ	0.022	0.027	0.104	0.213	0.416
Fix. duration NON-RA -> Loyalty	0.004	0.009	0.082	0.044	0.482
Fix. duration NON-RA -> Satisfaction	-0.030	-0.022	0.077	0.383	0.351

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values
Fix. duration NON-RA -> ThoughtsCQ	0.085	0.091	0.071	1.194	0.116
Fix. duration NON-RA -> ThoughtsOQ	0.013	0.008	0.098	0.132	0.448
Fixation NON-RA -> Loyalty	-0.002	-0.004	0.076	0.027	0.489
Fixation NON-RA -> Satisfaction	0.002	0.001	0.075	0.023	0.491
Fixation NON-RA -> ThoughtsCQ	-0.027	-0.029	0.075	0.357	0.361
Fixation NON-RA -> ThoughtsOQ	0.017	0.019	0.063	0.272	0.393
Fixation RA -> Loyalty	-0.066	-0.061	0.078	0.838	0.201
Fixation RA -> Satisfaction	-0.065	-0.060	0.073	0.894	0.186
Fixation RA -> ThoughtsCQ	-0.067	-0.067	0.072	0.936	0.175
Fixation RA -> ThoughtsOQ	0.003	0.001	0.082	0.039	0.484
Gender -> Loyalty	0.104	0.105	0.096	1.080	0.140
Gender -> Satisfaction	0.010	0.012	0.099	0.102	0.459
TfAQ -> Loyalty	-0.002	-0.001	0.054	0.035	0.486
TfAQ -> Satisfaction	-0.007	-0.006	0.052	0.143	0.443
ThoughtsCQ -> Loyalty	0.716	0.718	0.046	15.561	0.000
ThoughtsCQ -> Satisfaction	0.631	0.631	0.055	11.420	0.000
ThoughtsOQ -> Loyalty	0.046	0.044	0.050	0.923	0.178
ThoughtsOQ -> Satisfaction	0.160	0.160	0.047	3.397	0.000
Urgency -> Fix. Duration RA	0.392	0.394	0.204	1.926	0.027
Urgency -> Fix. duration NON-RA	0.429	0.436	0.197	2.179	0.015
Urgency -> Fixation NON-RA	0.096	0.102	0.183	0.524	0.300
Urgency -> Fixation RA	-0.269	-0.271	0.194	1.385	0.083
Urgency -> Loyalty	0.051	0.052	0.053	0.972	0.166
Urgency -> Satisfaction	0.033	0.037	0.047	0.709	0.239
Urgency -> ThoughtsCQ	0.046	0.051	0.055	0.828	0.204
Urgency -> ThoughtsOQ	0.015	0.009	0.060	0.249	0.402

	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values
Urgency x Confusion -> Fix. Duration RA	-0.604	-0.609	0.283	2.132	0.017
Urgency x Confusion -> Fix. duration NON-RA	-0.612	-0.622	0.292	2.096	0.018
Urgency x Confusion -> Fixation NON-RA	-0.150	-0.150	0.282	0.532	0.297
Urgency x Confusion -> Fixation RA	0.116	0.115	0.280	0.413	0.340
Urgency x Confusion -> Loyalty	-0.059	-0.062	0.067	0.886	0.188
Urgency x Confusion -> Satisfaction	-0.033	-0.037	0.060	0.553	0.290
Urgency x Confusion -> ThoughtsCQ	-0.046	-0.047	0.070	0.657	0.256
Urgency x Confusion -> ThoughtsOQ	-0.023	-0.010	0.081	0.292	0.385

Table 5.31 Specific indirect effects

Relationships	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values
Confusion -> Fix. duration NON-RA -> Loyalty	-0.022	-0.023	0.028	0.783	0.217
Confusion -> Fix. duration NON-RA -> Satisfaction	-0.033	-0.031	0.028	1.186	0.118
Confusion -> Fix. duration NON-RA -> ThoughtsCQ	0.033	0.035	0.035	0.936	0.175
Confusion -> Fix. duration NON-RA -> ThoughtsCQ -> Loyalty	0.023	0.025	0.025	0.932	0.176
Confusion -> Fix. duration NON-RA -> ThoughtsCQ -> Satisfaction	0.021	0.022	0.022	0.928	0.177
Confusion -> Fix. duration NON-RA -> ThoughtsOQ	0.005	0.002	0.041	0.119	0.453
Confusion -> Fix. duration NON-RA -> ThoughtsOQ -> Loyalty	0.000	-0.000	0.003	0.080	0.468
Confusion -> Fix. duration NON-RA -> ThoughtsOQ -> Satisfaction	0.001	0.000	0.007	0.113	0.455
Confusion -> Fix. Duration RA -> Loyalty	0.012	0.013	0.023	0.524	0.300
Confusion -> Fix. Duration RA -> Satisfaction	0.010	0.008	0.018	0.551	0.291
Confusion -> Fix. Duration RA -> ThoughtsCQ	-0.002	-0.001	0.022	0.098	0.461
Confusion -> Fix. Duration RA -> ThoughtsCQ -> Loyalty	-0.002	-0.001	0.016	0.097	0.461
Confusion -> Fix. Duration RA -> ThoughtsCQ -> Satisfaction	-0.001	-0.001	0.014	0.097	0.461
Confusion -> Fix. Duration RA -> ThoughtsOQ	0.003	0.005	0.025	0.115	0.454
Confusion -> Fix. Duration RA -> ThoughtsOQ -> Loyalty	0.000	0.000	0.002	0.073	0.471
Confusion -> Fix. Duration RA -> ThoughtsOQ -> Satisfaction	0.000	0.001	0.004	0.105	0.458
Confusion -> Fixation NON-RA -> Loyalty	0.002	0.003	0.012	0.124	0.451
Confusion -> Fixation NON-RA -> Satisfaction	0.002	0.004	0.012	0.122	0.451

Relationships	Original sample (O)	Sample mean (M)	Standard deviation (STDEV)	T statistics (|O/STDEV|)	P values
Confusion -> Fixation NON-RA -> ThoughtsCQ	-0.003	0.001	0.017	0.146	0.442
Confusion -> Fixation NON-RA -> ThoughtsCQ -> Loyalty	-0.002	0.001	0.012	0.146	0.442
Confusion -> Fixation NON-RA -> ThoughtsCQ -> Satisfaction	-0.002	0.001	0.011	0.145	0.442
Confusion -> Fixation NON-RA -> ThoughtsOQ	0.002	0.001	0.015	0.108	0.457
Confusion -> Fixation NON-RA -> ThoughtsOQ -> Loyalty	0.000	0.000	0.001	0.073	0.471
Confusion -> Fixation NON-RA -> ThoughtsOQ -> Satisfaction	0.000	0.000	0.003	0.102	0.459
Confusion -> Fixation RA -> Loyalty	-0.006	-0.003	0.021	0.273	0.392
Confusion -> Fixation RA -> Satisfaction	-0.007	-0.006	0.021	0.366	0.357
Confusion -> Fixation RA -> ThoughtsCQ	-0.022	-0.022	0.030	0.723	0.235
Confusion -> Fixation RA -> ThoughtsCQ -> Loyalty	-0.015	-0.016	0.022	0.715	0.237
Confusion -> Fixation RA -> ThoughtsCQ -> Satisfaction	-0.014	-0.014	0.019	0.728	0.233
Confusion -> Fixation RA -> ThoughtsOQ	0.001	-0.000	0.031	0.034	0.487
Confusion -> Fixation RA -> ThoughtsOQ -> Loyalty	0.000	-0.000	0.002	0.022	0.491
Confusion -> Fixation RA -> ThoughtsOQ -> Satisfaction	0.000	-0.000	0.005	0.032	0.487
Confusion -> ThoughtsCQ -> Loyalty	-0.549	-0.553	0.097	5.674	0.000
Confusion -> ThoughtsCQ -> Satisfaction	-0.483	-0.487	0.092	5.255	0.000
Confusion -> ThoughtsOQ -> Loyalty	-0.015	-0.015	0.019	0.769	0.221
Confusion -> ThoughtsOQ -> Satisfaction	-0.052	-0.052	0.028	1.866	0.031

Appendix 6: Hypotheses significance

Table 6.1: Hypothesis results

Hypothesis	Path	β	P	R_2
H1a	Customer confusion à satisfaction	- 0.535	< 0.001	0.622
H1b	Customer confusion à Loyalty	-0.364	< 0.05	0.586
H2	Customer confusion à Total fixation RA	0.322	< 0.05	0.048
	Customer confusion à Total fixation NON-RA	0.094	0.328	0.002
	Customer confusion à Average duration RA	0.128	0.244	0,032
	Customer confusion à Average duration NON-RA	0.382	< 0.05	0.029
H2a	Customer confusion x time urgency à Total fixation RA	0.116	0.340	0.048
	Customer confusion x time urgency à Total fixation NON-RA	-0.150	0.297	0.002
	Customer confusion x time urgency à Average duration RA	-0.604	< 0.05	0.032
	Customer confusion x time urgency à Average duration NON-RA	-0.612	< 0.05	0.029
H3a	Customer confusion à Total fixation RA à satisfaction	-0.007	0.357	0.622
	Customer confusion à Total fixation NON-RA à satisfaction	0.002	0.451	0.622
	Customer confusion à Average duration RA à satisfaction	0.010	0.291	0.622
	Customer confusion à Average duration NON-RA à satisfaction	-0.033	0.118	0.622
H3b	Customer confusion à Total fixation RA à loyalty	-0.006	0.392	0.586
	Customer confusion à Total fixation NON-RA à loyalty	0.000	0.451	0.586
	Customer confusion à Average duration RA à loyalty	0.012	0.300	0.586
	Customer confusion à Average duration NON-RA à loyalty	-0.022	0.217	0.586
H4	Confusion à Thoughts CQ	-0.760	<0.01	0.164

Hypothesis	Path	β	P	R_2
	Confusion à Thoughts OQ	-0.314	<0.05	0.028
H5a	Confusion à Thoughts CQ à satisfaction	-0.483	<0.001	0.622
	Confusion à Thoughts OQ à satisfaction	-0.052	<0.05	0.622
H5b	Confusion à Thoughts CQ à loyalty	-0.549	<0.001	0.586
	Confusion à Thoughts OQ à loyalty	-0.015	0.221	0.586
H6	Total fixation RA à Thoughts CQ	-0.067	0.175	0.164
	Total fixation NON-RA à Thoughts CQ	-0.027	0.361	0.164
	Average duration RA à Thoughts CQ	-0.017	0.428	0.164
	Average duration NON-RA à Thoughts CQ	0.085	0.116	0.164
	Total fixation RA à Thoughts OQ	0.003	0.484	0.028
	Total fixation NON-RAà Thoughts OQ	0.017	0.393	0.028
	Average duration RA à Thoughts OQ	0.022	0.416	0.028
	Average duration NON-RA à Thoughts OQ	0.013	0.448	0.028
H7a	Customer confusion à Total fixation RA à Thoughts CQ à satisfaction	-0.014	0.233	0.622
	Customer confusion à Total fixation NON-RA à Thoughts CQ à satisfaction	-0.002	0.442	0.622
	Customer confusion à Average duration RA à Thoughts CQ à satisfaction	-0.001	0.461	0.622
	Customer confusion à Average duration NON-RA à Thoughts CQ à satisfaction	0.021	0.177	0.622
	Customer confusion à Total fixation RA à Thoughts OQ à satisfaction	0.000	0.487	0.622
	Customer confusion à Total fixation NON-RA à Thoughts OQ à satisfaction	0.000	0.459	0.622
	Customer confusion à Average duration RA à Thoughts OQ à satisfaction	0.000	0.458	0.622
	Customer confusion à Average duration NON-RA à Thoughts OQ à satisfaction	0.001	0.455	0.622

H7b	Customer confusion à Total fixation RA à Thoughts CQ à loyalty	-0.015	0.237	0.586
	Customer confusion à Total fixation NON-RA à Thoughts CQ à loyalty	-0.002	0.442	0.586
	Customer confusion à Average duration RA à Thoughts CQ à loyalty	-0.002	0.461	0.586
	Customer confusion à Average duration NON-RA à Thoughts CQ à loyalty	0.023	0.176	0.586
	Customer confusion à Total fixation RA à Thoughts OQ à loyalty	0.000	0.491	0.586
	Customer confusion à Total fixation NON-RA à Thoughts OQ à loyalty	0.000	0.442	0.586
	Customer confusion à Average duration RA à Thoughts OQ à loyalty	0.000	0.471	0.586
	Customer confusion à Average duration NON-RA à Thoughts OQ à loyalty	0.000	0.468	0.586