Journal of Retailing and Consumer Services 57 (2020) 102217

Contents lists available at ScienceDirect

Journal of Retailing and Consumer Services

journal homepage: http://www.elsevier.com/locate/jretconser

A neurophysiological exploration of the dynamic nature of emotions during the customer experience

Nanouk Verhulst [a,b,c,*], Iris Vermeir [b], Hendrik Slabbinck [b], Bart Larivière [d,c], Maurizio Mauri [e,f], Vincenzo Russo [e]

[a] Vrije Universiteit Brussel (VUB), Department of Business, Pleinlaan 2, 1050, Brussels, Belgium
[b] Ghent University, Department of Marketing, Innovation and Organisation, Tweekerkenstraat 2, 9000, Gent, Belgium
[c] Center for Service Intelligence, Ghent University, Tweekerkenstraat 2, 9000, Gent, Belgium
[d] KULeuven, Department Marketing, Naamsestraat 69 – Box 3545, 3000, Leuven, Belgium
[e] IULM, Center of Research on Neuromarketing, Via Carlo Bo, 1, 20143, Milan, Italy
[f] SR Labs, Via Gaudenzio Fantoli, 7, 20138, Milan, Italy

ARTICLE INFO

Keywords:
Customer experience
Neurophysiological measurement
Emotions
Human versus machine

ABSTRACT

This paper explores the benefits of measuring emotions and their dynamic nature during the customer experience with neurophysiological measures. In this study, emotions are measured during a service interaction (with self-service technology or a human employee) going through a series of touchpoints, including a service failure. We show that creating a loyalty card with the help of a service employee or self-service technology did not impact customers' perceived service satisfaction and their behavioral intentions. This paper demonstrates that neurophysiological measures such as Galvanic Skin Response might be better equipped to unveil the dynamic nature of emotions (e.g., arousal) during the customer experience and that valence measured by neurophysiological tools (using electroencephalography) better reconciles with the effect found for satisfaction and behavioral intentions. Our findings have implications for both researchers and practitioners who want to understand and bolster customer experiences, thereby taking customer emotions and its appropriate measurement tools into consideration.

1. Introduction

Deeping the understanding of the customer experience (CX) has received notable attention in recent years (De Keyser et al., 2015; Lemon and Verhoef, 2016). Scholars recognize the key role of emotions during the entire CX (Gaur et al., 2014; Manthiou et al., 2020 this issue; Tronvoll, 2011), and the dynamic nature of both the CX and emotions (De Keyser et al., 2015, 2020; Kuppelwieser and Klaus, 2019). Understanding people's emotions (Bigné et al., 2008; Jani and Han, 2015) and how those emotions play a role during all stages and touchpoints of a CX is crucial (McColl-Kennedy et al., 2019; Puccinelli et al., 2009; Verhoef et al., 2009). Particularly, as emotional responses during the CX affect cognitive judgments of a service (Lerner and Keltner, 2000), service performance outcomes (e.g., satisfaction; Liljander and Strandvik, 1997), and behavioral intentions (e.g., return intentions; Ladhari, 2009). Despite its importance, studies on emotions during the CX typically relied solely on retrospective self-report measures (Benoit et al., 2017; Lajante et al., 2019), mostly assessing general emotions felt during the CX, hence studies exploring the dynamic nature of emotions during the CX remain scarce (Kuppelwieser and Klaus, 2019; Verhulst et al., 2019).

Extensive methods exist for measuring emotions during the CX (Klaus, 2014; Verleye, 2015), but most prior studies ignored the dynamic nature of emotions during the CX (Kuppelwieser and Klaus, 2019). Even though studies show that a dynamic measurement of the CX is of high importance (Verhoef et al., 2009). Prior studies primarily relied on self-report measures and scenario-based studies, which tend to restrict the internal and external validity of previous findings (Bruun and Ahm, 2015). Besides, the substantial reliance on retrospective self-reports scales also limited the opportunity to measure the dynamic

Abbreviations: CX, customer experience; SST, self-service technology.

* Corresponding author. Vrije Universiteit Brussel (VUB), Department of Business, Pleinlaan 2, 1050, Brussels, Belgium.
E-mail addresses: nanouk.verhulst@vub.be (N. Verhulst), iris.vermeir@ugent.be (I. Vermeir), Hendrik.Slabbinck@ugent.be (H. Slabbinck), bart.lariviere@kuleuven.be (B. Larivière), maurizio.mauri@srlabs.it (M. Mauri), vincenzo.russo@iulm.it (V. Russo).

https://doi.org/10.1016/j.jretconser.2020.102217

nature of emotions during the CX (Harrison-Walker, 2019; Verhulst et al., 2020). Indeed, self-report measures do not allow in-the-moment measurement of emotions and are suspectable to several types of measurement errors while measuring emotion (See Verhulst et al., 2019 for more details about measurement errors linked to self-report measures).

This paper uses neurophysiological measures to assess emotions along the CX, to advance the understanding of CX and its dynamics. The benefit of adopting neurophysiological measures to assess emotions is twofold. Neurophysiological measures allow capturing emotions in a real-time dynamic fashion during the CX and they can help to overcome several types of measurement errors (e.g., recall bias) associated with other types of measures (Lemon and Verhoef, 2016; Verhulst et al., 2019). Furthermore, understanding the implications of adopting neurophysiological measures versus conventional self-report measures is relevant because there is a need to validate old and newer methods to study the CX (Imhof and Klaus, 2019; McColl-Kennedy et al., 2015; Ostrom et al., 2015; Verhoef et al., 2009). Therefore, the *first research objective of this paper is to unveil the dynamic nature of the CX and explore the usefulness of neurophysiological tools (versus self-report measures) to capture emotions during the CX.*

To address the aforementioned research objective, we designed a real-life research setting conducted in a controlled lab. In this study, we consider a series of touchpoints including a service failure and a recovery to be able to examine the dynamic nature of experienced emotions. Including a failure was crucial, because it tends to spark negative emotions, which have a stronger impact on performance outcomes than positive emotions (Baumeister et al., 2001; Gijsenberg et al., 2015). By including both a failure and a recovery (associated with more positive emotions; e.g., Kuo and Wu, 2012), we can investigate if neurophysiological versus retrospective self-report measures are better equipped to pick up negative/positive emotions and/or emotional changes throughout the CX. Moreover, a better understanding of experienced emotions during service failure is key to improve behavioral predictions after such failures (Harrison-Walker, 2019).

Additionally, the impact of two service intervention types was examined, namely being served by a human (i.e., a service employee) versus a self-service technology. Specifically, the dynamic nature of the emotional components during the CX and its service performance outcomes were explored, while comparing service delivery by self-service technology (SST) versus human intervention.

We included the type of service interface (i.e., human versus machine) as it is an important antecedent of the emotional component of the CX (Verhoef et al., 2009). Furthermore, there is a renewed interest in studying both human and non-human frontline service interfaces and how those service interfaces can enhance the CX (Furrer et al., 2020; Singh et al., 2017). Specifically, technology recently transformed how services are delivered. Key in this transformation is the replacement of humans by machines ranging from self-service technologies to robotics (De Keyser et al., 2019; Larivière et al., 2017; Wirtz et al., 2018). In some types of services (e.g., E-retail, online banking), technology has completely replaced humans. However, the implications of such transitions on CX remain relatively unclear, especially when failures occur, and customers might prefer a human intervention to fix the failure (De Keyser et al., 2015). The impact of failure during service interactions with human employees is well understood, nevertheless how this translates to service interactions with SST remains unclear (Köcher and Paluch, 2019). For instance, the SST literature to date mainly focused on determinants of customers' intentions to adopt and using SST, such as technological readiness and ease of use (e.g., Blut et al., 2016; Dabholkar and Bagozzi, 2002; Lin and Hsieh, 2007; Oh et al., 2013; Wang et al., 2013). Nevertheless, little is known about emotions during service delivery by machine versus a human employee, how those emotions fluctuate throughout a CX, and their impact on service performance outcomes (e.g, Köcher and Paluch, 2019). As a result, our *second research objective is to explore which specific emotional components (i.e. valence and arousal) are experienced during different touchpoints of the CX while a service is delivered by SST or a human employee.* Specifically, this paper explores if the service interface-type (human intervention versus SST) affects emotions differently in the various phases throughout the customer journey and the impact of the type of service interface on service performance outcomes including customer satisfaction and behavioral intentions.

This paper contributes to the literature in three important ways: (1) we demonstrate that neurophysiological measures can be better equipped to unveil the dynamic nature of emotions during the CX. Particularly, when the objective is to understand the customer's level of *arousal* over time (i.e., Research objective 1), (2) valence measured in-the-moment by neurophysiological measures was not found to be different between human intervention versus SST encounters. This better reconciles with the found effect on the outcome variables (i.e., customer satisfaction and behavioral intentions), and outperforms the self-report measures and observed relationships (i.e., Research objective 2) and (3) our results reveal that self-report *valence* measures can be biased and overrated when customers recall failure and touchpoints that require additional customer efforts caused by these failures.

In the following sections, we briefly discuss retrospective versus neurophysiological measures, SST literature, and compare services delivered via humans versus machines. Next, we discuss the study design, methods, and results. We then elaborate on theoretical, methodological, and managerial implications, before discussing limitations and future research ideas.

2. Theoretical background

2.1. Measuring emotions by neurophysiological versus self-report measures

Previous service research on the CX and emotions has predominantly relied on retrospective self-report measures. Adopting self-reports studies are generally cheap, can offer rich data, allow fast data-collection, and straightforward data analyses, yet they do have several downsides especially while measuring emotions along the CX (Hulland et al., 2018; Podsakoff and Organ, 1986; Verhulst et al., 2019). First of all, retrospective self-reports do not allow in-the-moment dynamic measurement of a CX (Boshoff, 2012; Kahneman and Krueger, 2006; Verhoef et al., 2009). Hence, retrospectively reported emotions can be based on representations about emotions rather than on emotions itself, which could result in different results (Clore et al., 2001; Robinson and Clore, 2001). Thinking about how you felt in the past (retrospective feelings) does not necessarily reflect how you really felt at that time or people may not accurately recall experienced emotions (e.g., recall bias; Pizzi et al., 2015). Further, while retrospectively reconstructing emotions, actually experienced emotions and contextual details are absent (Lazarus, 1995). Not remembering or not reporting actual emotions does not have to be a problem. In some contexts, a retrospective representation of emotion can be accurate and might be enough to predict consumer outcomes or behavior (Harmon-Jones et al., 2016), but research on comparing the representation of emotions (e.g., self-report) versus measuring actual experienced emotions (e.g., neurophysiological tools) to predict service outcomes and consumer behavior is still in a nascent state (e.g., Boksem and Smidts, 2014) and may reveal novel insights into emotions, CX and (consumer) behaviors.

Secondly, self-assessment of emotions experienced during CX can be biased. For instance, people might exaggerate answers, be unwilling to answer, answer in a socially desirable way, or give overly positive responses since the emotions or outcomes reported are already processed by their mind (Clore et al., 2001; Podsakoff et al., 2003; Poels and Dewitte, 2006).

In contrast to retrospective self-reports, neurophysiological measures allow capturing temporal fluctuations of emotions during a CX (Boshoff, 2012; Liapis et al., 2015). Additionally, they allow measurement of both automatic and (sub)conscious emotional responses during the CX in real-time (instead of post-experience). Thus, neurophysiological measurements not only permit to identify key moments and capture the dynamic nature of the CX (Zomerdijk and Voss, 2010), but are more likely to resemble true feelings associated with actual experiences (Kahneman and Krueger, 2006). Applying neurophysiological tools can enrich the overall understanding of affective processes and can side-step several sources of measurement error linked to self-report (Bell et al., 2018; Lajante and Ladhari, 2019). For example, item-nonresponse, issues with response styles, order effects, and social desirability are not present, since participants do not have to answer questions, instead, automatic bodily changes (linked to emotion) are measured in real-time (Verhulst et al., 2019). To illustrate, Bruun and Ahm (2015) showed an inconsistency between retrospective emotion ratings and the in-the-moment measures during human-technology interaction. Retrospective ratings (versus in-the-moment) overestimated negative emotions during unpleasant interactions, whereas during pleasurable interactions retrospective and in-the-moment measures better reconciled.

Neurophysiological measures are very suitable to explore affective processes throughout a service interaction (Bell et al., 2018; Lajante and Ladhari, 2019; Verhulst et al., 2020), but the application of these tools is relatively expensive (equipment, payment participants), time-consuming (e.g., install sensors, data analysis), and data analysis is more complex, such that managers and scholars might decline it (Verhulst et al., 2019). Recently, researchers have suggested that combining neurophysiological measures with more traditional measures (e.g., self-reports) can enhance our field (Bell et al., 2018; Lajante and Ladhari, 2019; Verhulst et al., 2019). Not only because both types of measurements can potentially add unique information to topics under investigation, but also because they can overcome different types of limitations. Hence, the service literature needs to explore if it is worthwhile to adopt neurophysiological measures to study the CX, since adding them to your toolkit comes at a cost.

This paper uses two neurophysiological tools, namely electroencephalography (EEG) and sensors to measure galvanic skin response (GSR) to measure emotions. Adopting both EEG and GSR has several advantages. Firstly, they both measure a different dimension of emotions, namely EEG is, amongst other things, related to valence (or pleasure; Harris et al., 2018; Ohme et al., 2009) and GSR to arousal (Christopoulos et al., 2019). EEG and GSR are well-established and validated measures of valence and arousal. Secondly, they are both highly suitable to capture in-the-moment temporal fluctuation (Harris et al., 2018; Kenning et al., 2007). Furthermore, they are not susceptible to biases such as recall bias (e.g., Bell et al., 2018; Verhulst et al., 2019). Lastly, they are both relatively affordable, not very invasive, and not overly complex to use and interpret, compared to other neuroscientific tools (Caruelle et al., 2019; Harris et al., 2018).

2.2. Emotions while being served by self-service technology or a human

SSTs are defined as *"technological interfaces that enable customers to produce a service independent of direct service employee involvement"* (Meuter et al., 2000, p. 50). Examples include self-scanning checkouts at retailers and self-check-in at airports. SST research to date has mainly focused on determinants of customers' intentions to adopt and use SST, such as technological readiness and ease of use (e.g., Blut et al., 2016; Dabholkar and Bagozzi, 2002; Lin and Hsieh, 2007; Oh et al., 2013). Yet, less is known about emotions during SST usage, how emotions fluctuate throughout an SST experience, and the differential effects of service interfaces (SST versus human intervention) on various phases of a CX (e.g., Vakulenko et al., 2019).

Past research on the effects of employee-customer (i.e., human intervention) suggests that interpersonal interactions during service delivery have a substantial effect on overall service performance outcomes, with good interactions improving the CX (Bitner et al., 1990). Indeed, frontline employees have an enormous impact on customers' emotions (for an overview of relevant studies see Petzer et al., 2012).

Similar processes are at play during service encounters with SST or a human employee. However, the discussion on which of the two is preferred continues to be an important subject of debate (Köcher and Paluch, 2019). Further, many people believe, without strong or converging empirical back-up, that SST is less satisfying than human service delivery. This rather negative perception of SST exists in both academia (Shi, 2011; Yan et al., 2013) and business (Anand, 2011). But in reality, only a few studies have empirically investigated the different impact of service delivery by SST versus an employee on service performance outcomes.

Several researchers suggest that people generally prefer to be served by a human employee (versus SST) and that SST is less satisfying than human intervention. For instance, Kattara and El-Said (2014) tested preferences for SST versus human intervention for different service encounters during a hotel stay. They found that for most types of encounters (e.g., check-in, reservation) customers prefer human intervention over a technology-based interaction. Yan et al. (2013) appear to be the first to attempt to understand how SST versus human service delivery impacts customers' outcomes within the service domain. They showed that consumers perceived higher levels of service quality after service delivery by an employee versus SST, which in turn led to higher satisfaction. Other studies suggest a preference for human service delivery over SST by showing that including human features or higher perceived warmth in SST makes people less likely to switch back to service delivery by an employee, and leads to better attitudes about SST and improved service outcomes (Appel et al., 2012; Fan et al., 2016; van Doorn et al., 2017).

On the other side, some scholars do not fully agree that human intervention is always preferred or that higher satisfaction arises from human intervention (Espina and Pérez, 2014; Ha and Perks, 2005; Janda and Ybarra, 2005; Köcher and Paluch, 2019; Shankar et al., 2003). SST may have specific features that induce a positive CX (e.g., shorter perceived waiting time, increased perceived control; Dabholkar et al., 2003; Djelassi et al., 2018; Weijters et al., 2007). Note that also service context (e.g., tourism, banking) might well be a defining factor related to preferences and service outcomes related to human service delivery versus SST (Edvardsson et al., 2018).

Next to the mixed findings for human service delivery versus SST on service performance outcomes, differential effects of service interfaces (SST versus human service delivery) on various phases of a CX remain unclear. This is surprising since it is generally known that consumers' emotional responses change throughout a CX (e.g., Boshoff, 2012; Liu et al., 2016). Exceptions are the studies of Mattila et al. (2011) and Harris et al. (2006) who suggest that recovery actions are better conceived or more satisfying when delivered by an employee instead of SST. This implies that emotional reactions towards the recovery phase of CX should be more positive when delivered by an employee versus SST. However, they let consumers judge scenarios using retrospective self-reports which have their limitations, and they ignored other phases during a CX. Recently, Köcher and Paluch (2019) showed a tendency for higher service satisfaction after a service interaction with SST versus human intervention. They also showed that re-use intentions remain similar during SST interaction with or without a failure, whereas human service delivery with a service failure (versus no failure) resulted in lower return intentions.

In sum, previous literature is not entirely conclusive about whether human service delivery outperforms SSTs. Several reasons exist for these mixed findings, and our study attempts to overcome some of these issues. First, ample research has been conducted on SST and human intervention in isolation, but empirical evidence directly comparing emotions and performance outcomes following SST use or human service delivery is scarce. Second, a wide variety of designs and measurement tools such as experiments, field studies, and questionnaires have been used to investigate service delivery by employees and SST, which can result in different findings. Most research has relied on post-encounter self-report measures, which have several pitfalls, as previously explained. Third, most studies have used scenarios rather than real interactions, which could result in participants making judgments based on their general preferences or common beliefs instead of what they would truly experience (Harrison-Walker, 2019). Finally, previous research did not account for the dynamic nature of emotions during the CX. For example, customers' emotions can vary during a failure or recovery phase and can affect service performance outcomes differently (Harrison-Walker, 2019). To tackle these shortcomings, our study directly compares the emotional impact of service delivery by SST and human intervention. Further, we measure emotions by a multi-method approach that allows accounting for the dynamic nature of both conscious and subconscious emotional components of CX and overcoming issues with post-encounter self-report measures. To overcome limitations of scenario-based research, we used a real-life interaction in which participants really talked to an employee or really used an SST (in contrast to "reading" a scenario and asking respondents imagining themselves in that scenario).

3. Methodology

3.1. Research design and study overview

First, a pilot test (N = 4) was carried out to ensure that all neurophysiological tools were working properly and that the flow of the design worked well. Next, we invited 40 participants (50% men, M_{age} = 23 years), including students and personnel of a European University, to our laboratory to participate in a 2 (SST versus human intervention, between subject) × 6 (phases, within subjects) mixed design experiment.

3.1.1. Conditions

Participants in our study made a loyalty card either by using SST (SST condition) or by communicating with a real-life employee (human condition) while neurophysiological responses were measured.

Both conditions were the same except that during the human condition an employee (that was really present) requested the participants' personal information and typed the information on a computer for them. Whereas in the SST condition participants read instructions on the personal information they had to provide and typed this information in themselves on a computer. What the employee said to the participants during the human condition was exactly the same as what participants saw on the screen of the SST (SST condition). Furthermore, when participants had to type information in the SST system (SST condition), the employee typed the information provided by participants in a computer (cf. phases see below). Participants were randomly assigned to one of the experimental conditions. However, we did balance the gender distribution between conditions, because men and women differ in certain biological responses, such as galvanic skin response (GSR; Bianchin and Angrilli, 2012). In both conditions, participants went through several phases (e.g., failure) during this real-life service encounter.

3.1.2. Service phases to capture emotional dynamics

The service encounter consisted of several phases (cf. Boshoff, 2012; see Fig. 1). See Appendix A for the exact sentences presented by the SST or said by the human employee (i.e., the detailed script). First, participants were greeted during the welcome phase. In the human condition, an employee speaks to welcome the participants, whereas during the SST condition the participants read the welcome text on the screen. Next, they either typed their personal details (e.g., name; SST condition) or gave the employee their details (Service intervention 1 phase; human

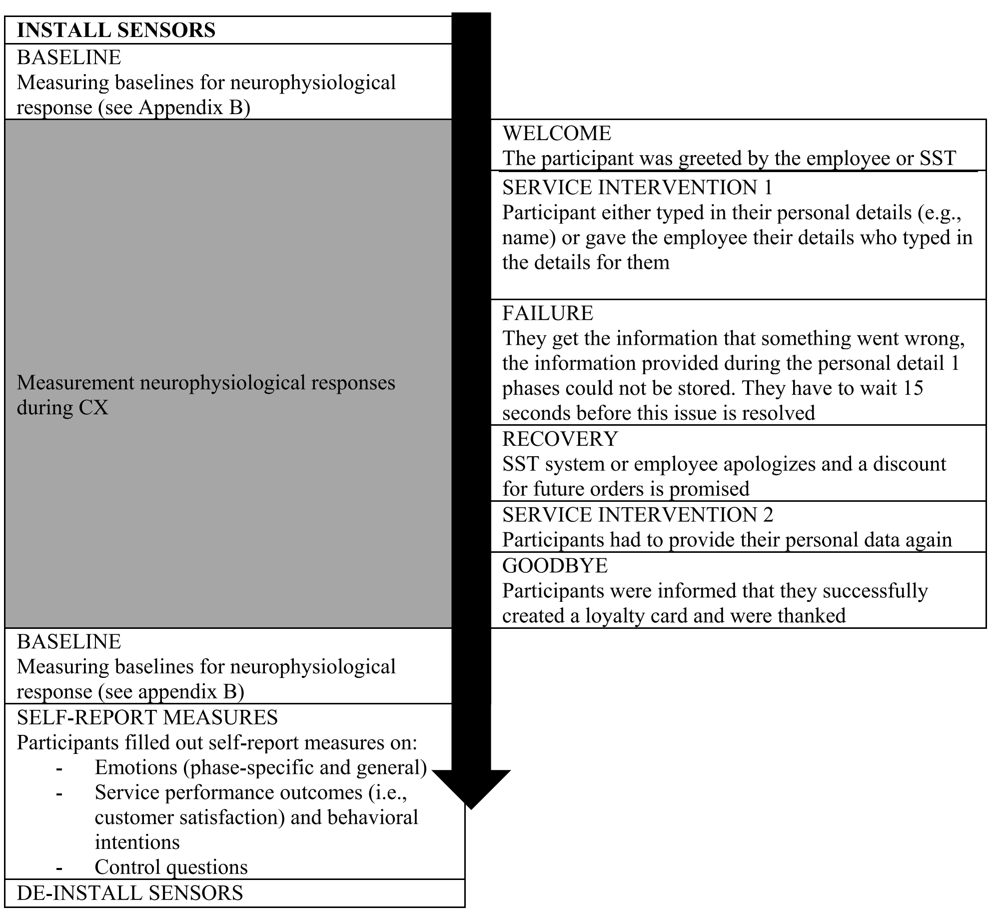

Fig. 1. Order of the design and measurement.

condition). After giving their details; a failure happened (participants read or the employee said: "Something went wrong, we could not store your information. Please wait a moment"; failure phase). The duration of the failure was 15 s. Importantly, the failure was independent of the service interface (i.e., human or SST condition). In both conditions, the computer system (i.e., the SST or the computer used by the employee) could not store the participant's information, preventing participants from having different blame attributions between conditions, which otherwise might result in different behavior (Harris et al., 2006). After these 15 s, a recovery action was undertaken (excuses and promised discount for future orders; recovery phase; both in SST and Human condition the same excuses and promised discount were provided). Afterward, either the employee or the SST-system asked participants to provide their data again (Service intervention 2 phase). Finally, they were informed that they successfully created a loyalty card and were thanked (goodbye phase).

The rationale to include various phases during this service encounter was mainly to be able to examine the dynamic nature of emotions. If we would not have included failure and recovery it would be hard to investigate the emotional changes throughout a CX. Furthermore, a welcome and goodbye are a standard part of most types of service interaction, thus important to be included to not deviate from minimum service expectations (e.g., Solomon et al., 1985). Furthermore, service intervention 1 and 2 had to be separate phases for methodological reasons, namely different baseline measures are necessary to correct the brain activity and GSR responses during these two phases, since participants had to either type or talk during these phases which can activate changes in GSR or brain activation in itself (e.g., Bell et al., 2018). Additionally, to make the experiment as realistic as possible we included service intervention 2, where participants had to repeat/re-type the information that could not be stored during the failure phase.

3.1.3. Procedure

First, participants entered the laboratory, read an instruction letter explaining the neurophysiological tools and study purpose, and signed an informed consent document. Respondents were told that their neurophysiological responses would be measured while creating a bookshop loyalty card to test the "service script" of a specific company. This cover story was identical between conditions. Next, participants were seated in a comfortable chair while EEG and GSR sensors were placed. The experiment began with several steps: baseline measurements, the service encounter, and filling out a survey (see Fig. 1). In line with standard practices (e.g., Verhulst et al., 2019), we measured brain activity (by EEG) and GSR before starting the real experiment to establish baselines for each participant. The default brain activity and skin conductance are different for each individual (e.g., Caruelle et al., 2019; Jackson et al., 2003), hence baseline should be used to correct experimental responses (Kirk, 2003; Zhang et al., 2014). For a detailed discussion of baseline procedure and baseline calculations see Appendix B.

The actual experiment started upon completion of the baseline measurements. The storyline of the encounter was the same in both conditions (see Appendix A), but in the SST condition participants were welcomed by a computer and had to create a loyalty card themselves by using SST. In the human condition participants interacted with a service employee (i.e., research assistant). The research assistant had undergone training to always act in exactly the same manner. In addition, he was always dressed in the same clothing because the appearance of service employees can influence service perceptions (Shao et al., 2004).

After the participants went through all six phases of the service encounter, baselines were again measured (see appendix B for baseline procedure). After this procedure, a laboratory assistant entered the room and stopped the recording of neuropsychological measures and opened the retrospective self-report survey on emotions and performance outcomes. Once the survey was finished, the neuropsychological tools were disconnected. Lastly, before leaving the laboratory, participants were thanked for their involvement with an invitation to participate in a raffle to win a 30-euro coupon for a multi-media store.

3.2. Measuring emotions, their dynamic, and outcomes

To investigate emotions, we draw from Stimulus-Organism-Response models (e.g. (social) servicescape model; Bitner, 1992; Tombs and McColl-Kennedy, 2003), which state that environmental factors, such as service interface, influence customers' emotions. Emotions, in turn, affect attitudes toward, for instance, service providers. Emotions generally are comprised of two dimensions, namely valence (or pleasure) and arousal (cf., circumplex model of emotion; Russell, 1980). Emotion (valence and arousal) was measured both by neurophysiological (GSR, EEG) and retrospective self-report methods.

3.2.1. Neuropsychological measures

We used two neurophysiological measures (for more details on these measurements see Appendix C). Changes in GSR were recorded through electrodes placed on the shoulder (see Fig. 2). These changes are used to measure stress reactions, and they can indicate emotional arousal (Caruelle et al., 2019; Christopoulos et al., 2019).

Frontal brain activity was monitored using two EEG-sensors on the scalp (see Fig. 2; Tivadar and Murray, 2019). We assessed frontal brain asymmetry which indicates approach-avoidance behavior during the service experience (Mauss and Robinson, 2009; Ohme et al., 2011). Approach behavior is associated with positive emotions, such as engagement, interest, and happiness (Davidson et al., 1990). Avoidance, on the other hand, is associated with negative emotions, such as disinterest, disengagement, fear, and disgust (Davidson et al., 1990). Consequently, asymmetry is a promising measure of emotional valence (Ohme et al., 2009). The asymmetry score was calculated by subtracting the left alpha power from right alpha power (cf. standard practices; Coan and Allen, 2003). Positive scores mean relatively more approach behavior than avoidance behavior, whereas negative scores mean relatively more avoidance behavior.

3.2.2. Self-report measures

Most measures were adopted from past research, and all items were translated from English into the native language of the research facility by means of back-translation (see Appendix D for items, items order, and Cronbach's alphas). Most measures used a 7-point Likert scale and had satisfying Cronbach's alphas ($\alpha > 0.6$).

3.2.2.1. Emotions. Participants answered items about both their general emotions (i.e., global valence and arousal) related to the service experience and their phase-specific valence. For general emotions, participants first indicated how they felt during this service encounter (Ladhari, 2009; ranging from not at all pleasant to pleasant). Followed by the 12-item semantic differential scale of Russell (1980) to assess valence (e.g., unhappy to happy) and arousal (e.g., cheerful to depressed). The items "active -passive", "cheerful-depressed", "surprised-indifferent" were deleted from the arousal scale, because a confirmatory factor analysis (CFA) revealed that the factor loadings for these three items were insufficient (all lower than 0.5). More detailed information about the results of the CFA is reported in Appendix D.

For phase-specific valence, participants indicated if they experienced negative, neutral, or positive feelings (not at all to very much) during all six phases separately (e.g., greeting). To ensure that participants knew which phase we were referring to, the text that was used during the service experience was shown at the top of the answer page.

3.2.2.2. Service performance and behavioral intentions. Participants rated three satisfaction items (Lin and Hsieh, 2007; Maxham, 2001) as a measure for their perceived service performance. In addition, the survey included three types of behavioral intentions: three items about

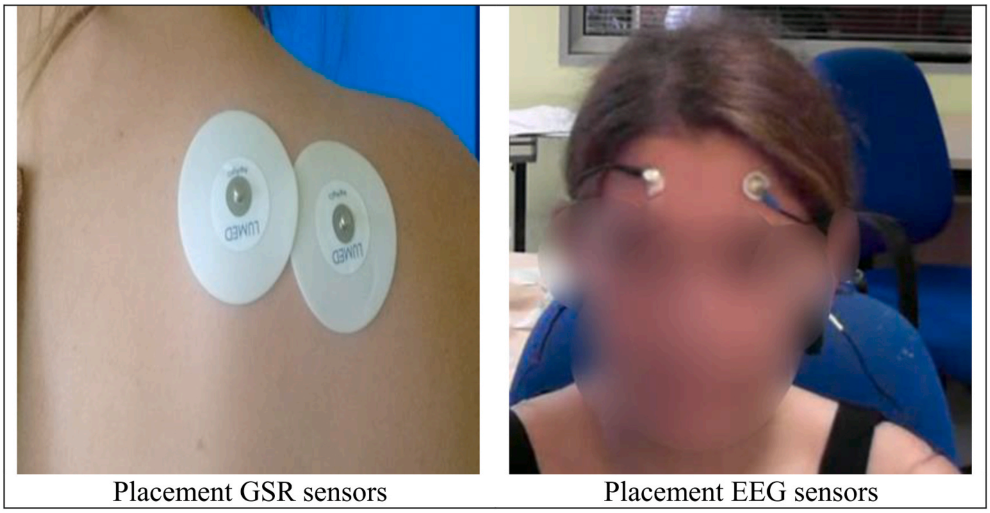

Fig. 2. Placement of GSR and EEG sensors.

purchase intent in the bookstore (Maxham, 2001), three re-use intention items (Turel et al., 2008), and one complaint intention item (Kim et al., 2003).

3.2.2.3. Control questions and scenario checks. Participants answered four items about their need for human interaction (Dabholkar, 1996) and two items about their habit of using SST (self-constructed). Their gender and age were also assessed. Gender, need for human interaction, and habit of using SST were used as covariates in our data-analysis (Blut et al., 2016; Venkatesh et al., 2012). As a check for the scenarios (or scripts) used, participants filled out one item about scenario realism, one for scenario valence, two for ease of use to create a loyalty card (self-constructed), and three about their perceived control while creating the loyalty card (adapted from Dabholkar, 1996). These scenario checks were used to ensure that the scripts were equal for those variables in both conditions since they can influence reactions of consumers toward an SST service. For example, past research has suggested that using SST might increase feelings of perceived control (Collier and Sherrell, 2010; Robertson et al., 2016), which could impact customer satisfaction. Furthermore, ease of use can also be an important antecedent of satisfaction (Robertson et al., 2016).

4. Results

4.1. Scenario checks

An independent *t*-test revealed that realism was sufficient (M_{SST} = 5.65; M_H = 5.45) for both conditions and did not significantly differ ($t(38) = -0.37, p > .05$). This indicates that both scenarios were perceived as equally realistic. Furthermore, scenario valence (e.g., positive or negative perceived service) was the same in both conditions ($t(38) = 0.48$, $p > .05$; $M_{SST} = 4.70$; $M_{Human} = 4.55$). Additionally, perceived control ($t(38) = -0.07, p > .05$; $M_{SST} = 4.03$; $M_{Human} = 4.07$) and ease of creating the loyalty card ($t(38) = -0.32, p > .05$; $M_{SST} = 4.60$; $M_{Human} =$ 4.75) did not significantly differ between conditions. For all the following analyses we used gender, need for human interaction, and habit of using SST as control variables, and previously described baselines were also used as covariates. For means and standard deviations for the scenario checks and control questions see Table 1.

4.2. Effect of service interface on emotions

We first conducted a series of 2 (SST versus human intervention, between subjects) × 6 (different phases, within subjects) mixed linear model analyses to investigate the effects of service interface on emotions and more specifically, during different phases. We used the first-order autoregressive model (ARH(1)) as repeated covariance type to account for the repeated nature of the data (i.e., emotions experienced in an earlier phase may impact the experienced emotions in a later phase, but not the other way around). Indeed, neurological data from a certain phase have an effect on the following phase. To control for individual differences in GSR and brain activity, we used the baselines as covariates in our models (Pocock et al., 2002; Zhang et al., 2014). Remember, the service encounter had six phases, namely the welcome phase, service intervention 1 (proving personal details), a failure (something went wrong with the data storage and a 15-s wait), recovery phase (an apology and a promised discount for future orders), service intervention 2 (providing their personal information again), and a goodbye phase (see 3.1.2, Fig. 1, and Appendix B for details).

First, results for arousal measured by GSR ($arousal_{GSR}$) and valence by EEG ($valence_{EEG}$) are discussed, which are followed by the results from the self-reports. We only found a significant main effect for phase ($F(5, 157.46) = 2.64$, $p < .05$; $\eta 2_p = .07$) on $arousal_{GSR}$ (see Table 2 for estimated means). No significant effect of service interface on arousal was found. To follow up on the main effect of phase, we conducted pairwise comparisons of the different phases with Bonferroni correction. Arousal during the goodbye phase was significantly different from the recovery phase (with $p < .05$) and marginally significantly from the service intervention 1 phase and welcome phase (with $p < .10$). Arousal was highest during the goodbye phase and increased throughout the whole service encounter. Further, except for baseline values, none of the control variables had an effect on $arousal_{GSR}$.

For $valence_{EEG}$ we did not find any significant effects of phase or service interface ($p > .05$; see Table 2 for estimated means). Again, none of the control variables had an effect on $valence_{EEG}$, except baseline values. Note that positive $valence_{EEG}$ scores indicate relatively more

Table 1
Means and standard deviations for scenario checks and control variables.

	SST	Human
Scenario checks		
Realism	5.65 (1.66)	5.45 (1.73)
Valence	4.70 (1.13)	4.55 (.83)
Perceived control	4.03 (1.61)	4.07 (1.22)
Ease of use	5.60 (1.62)	5.75 (1.29)
Control questions		
Habit of using SST	5.40 (1.57)	5.35 (1.63)
Need for human interaction	4.48 (1.06)	5.13 (1.23)

Table 2
Estimated means for $arousal_{GSR}$ and $valence_{EEG}$ (measured by neurophysiological measures).

	$Arousal_{GSR}$	$Valence_{EEG}$
Service interface		
SST	11.97	.88
Human	11.89	.09
Phases		
Welcome	11.65	.669
Service intervention 1	11.71	.230
Failure	11.90	.545
Recovery	11.90	.844
Service intervention 2	12.08	-.286
Goodbye	12.34	.896

approach behavior (versus avoidance behavior), whereas negative scores mean relatively more avoidance behavior (Coan and Allen, 2003).

To investigate self-reported emotions, we first conducted a 2 (SST versus human intervention, between subjects) × 6 (different phases, within subjects) mixed linear model analysis with phase-specific valence ($valence_{phase}$) as the dependent variable. A diagonal repeated covariance type was used. We found a main effect of phase ($F(5, 61.04) = 28.61$, $p < .001$; $\eta 2_p = .70$) and a marginally significant effect of service interface on $valence_{phase}$ ($F(1, 217.60) = 3.13$, $p < .10$; $\eta 2_p = .01$; see Table 3 for estimated means). $Valence_{phase}$ was more positive in the human condition than in the SST condition. The interaction between service interface and phase on $valence_{phase}$ was not significant ($p > .05$; $\eta 2_p = .05$). Following up on the main effect of phase, we conducted pairwise comparisons of the different phases with Bonferroni correction (see Fig. 3). Both the failure and filling out service intervention 2 phase significantly differed from all other phases (with $p < .001$). The other phases did not differ significantly from each other. $Valence_{phase}$ was most positive during welcome, service intervention 1, recovery, and goodbye phases, whereas the most negative during failure and service intervention 2 phase. Regarding the control question, both gender ($M_{man} = 4.10$, $M_{female} = 4.54$; $F(1, 215.44) = 11.15$, $p < .001$) and need for human interaction ($F(1, 215.44) = 7.40$, $p < .05$) had a significant impact on $valence_{phase}$. A higher need for human interaction leads to higher $valence_{phase}$.

To investigate the effect of service interface on self-reported global valence and arousal, we conducted two ANOVAs. We found no significant main effects of condition on global valence ($F(1, 35) = 1.31$, $p > .05$; $\eta 2_p = .04$) and arousal ($F(1, 35) = 0.34$, $p > .05$; $\eta 2_p = .01$). Furthermore, regarding the control variables only need for human interaction had an influence on arousal ($F(1, 35) = 4.11$, $p = .05$). Additionally, gender ($M_{man} = 4.10$, $M_{female} = 4.85$; $F(1, 35) = 9.14$, $p < .05$) and habit of using SST ($F(1, 35) = 9.14$, $p < .10$) had (marginally) significant effects on global valence.

In sum, when looking at the neurophysiological data, we only found an effect of phase on arousal. When looking at the self-reported retrospective global emotions, we do not find any effect of service interface

Table 3
Estimated means for phase-specific valence via self-report measures.

	Phase-specific Valence across phases by service interface
SST	4.20
Human	4.44
	Phase-specific Valence per phase
Welcome	4.80
Service intervention 1	4.53
Failure	2.98
Recovery	5.15
Service intervention 2	3.45
Goodbye	5.00

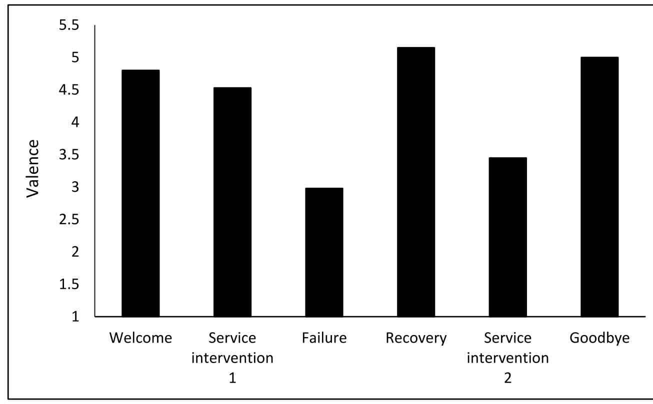

Fig. 3. Main effect of phase on self-reported phase-specific valence.

on valence and arousal. However, when we consider self-reported phase-specific valence, we see that phase has a clear impact and service interface a marginally significant impact. The impact of service interface on emotions seems very limited in our study. However, emotions during the different phases do clearly fluctuate throughout the CX.

4.3. Effect of service interface on service performance evaluation outcomes

To investigate the effect of service interface on customer service evaluations, we conducted a series of ANOVAs. We found no significant main effect of service interface on satisfaction, purchase intention, complaint intention, and re-use intention (with $p > .05$). When looking at the control questions, we see several effects. For instance, gender (marginally) significantly influenced satisfaction, purchase intention, and re-use intention. Further, the habit of using SST affected satisfaction. In sum, we found no effects of service interface on the service performance outcomes.

5. General discussion

5.1. Summary

This paper explored whether measuring emotions with neurophysiological measures adds to our understanding of the dynamic nature of the CX (Research objective 1). This paper shows that neurophysiological measures can help to gain insight into the dynamic nature of emotions during the CX. Further, we show that arousal was present (and increasing over time), although not self-reported by participants. When certain emotions are not stated in self-reports this does not mean they are not present, should be ignored, or have no impact (Harmon-Jones et al., 2016; Verhulst et al., 2019). Hence, measuring arousal with GSR can make more sense, since self-reported arousal might just pick up the most intense arousal events (e.g., linked to anger; Myrtek and Brügner, 1996). This is especially the case for people that have low self-awareness of bodily cues linked to emotion (Barrett et al., 2007; Salgado and Kingo, 2019). Furthermore, we aimed to investigate the usefulness of self-report measures versus neurophysiological measures while studying the CX. Furthermore, we clearly observed differences between the two methods. It seems that participants report more phase-specific emotions (valence) than actually experienced, thus there is a certain danger to measuring emotions consciously by means of self-reported tools. In this study, findings obtained by the neurophysiological measures better reconciled with the observed outcomes in this study.

The second aim of this paper was to explore the emotional component of the CX while a service is delivered by SST or an employee (Research objective 2). We show that in the context of creating a loyalty card SST service delivery is as good as human service delivery and that neurophysiological measures were needed to better understand this finding. Overall, in testing customer reactions to technology versus humans, and even to understand failures, neurophysiological tools might offer valuable insights – even if they show no effects (in contrast to self-reported emotions).

5.2. Discussion of findings

Table 4 provides an overview of the main findings of this study, which are discussed in the next paragraphs. Our results show that service interfaces and phase have no effect on $valence_{EEG}$, but we did find a main effect of phase and a marginally significant effect of service interface on self-reported phase-specific valence. The self-reported $valence_{phase}$ was more positive during the human condition. Further, in line with what would be expected (e.g., Boshoff, 2012), the failure and filling out service intervention 2 phase were the most negative and differed from all other phases. That the failure phase would be the most negative was proposed in most past research, except for Boshoff (2012). For the service intervention 2 phase, we argue that people are probably a bit annoyed that they must give their details again even though they received a discount. Important to note is that, in line with Boshoff (2012), but in contrast to many other researchers, we did not find higher positive effects (compared to other phases) of recovery actions (e.g., Chebat and Slusarczyk, 2005; DeWitt et al., 2008). We also did not find that recovery actions are better perceived when done by a human employee, contrasting results of, for instance, Mattila et al. (2011).

In addition, we did not find an effect of service interface on self-reported global valence. One may conclude that service interface essentially has no effect or only a limited effect on valence while creating a loyalty card. We suggest that when people are confronted with post-encounter valence questions about specific phases (i.e., obliged to form an opinion), they are more likely to report negative valence, for instance, for the failure phase. It is possible that these answers do not reflect their true feelings during the encounter. Instead, they could stem from a rather forced post-encounter cognitive appraisal (Cacioppo, 2002) or customers' answers can become more extreme in this case, especially with negative emotions (e.g., Bruun and Ahm, 2015). On the other hand, unlike neurophysiological measures, a post-encounter global valence measure cannot detect the subtlety of emotions varying throughout a CX.

Furthermore, looking at the neurophysiological data, we found a main effect of phase on arousal. Interestingly, no effect was observed on self-reported arousal. A plausible explanation, in line with peak-end theory, is that people are only able to recall their general level of arousal, arousal they experienced during an event, or at the end of an experience (Kahneman et al., 1997; Verhoef et al., 2004). For instance, if they get especially aroused during one phase, they may remember this best and report this arousal. This situation illustrates the advantages of continuous and in-the-moment measurement of arousal by means of GSR during a CX. Capturing the intensity of emotion throughout a CX in combination with emotional valence can offer new insights into understanding customer behavior better (Caruelle et al., 2019; Harrison-Walker, 2019).

Contrary to previous literature, we did not find differences of service interface on service performance outcomes such as satisfaction and re-use intention. This indicates that during this context (i.e., when creating a loyalty card) it does not matter if customers use SST or are assisted by an employee. This finding might be explained by the service interface and the context of our study. Specifically, creating a loyalty card is a simple service encounter. Hence, this type of service encounter – albeit a failure happened during this customer journey of creating a loyalty card – might be responsible for the limited impact we observed for the self-reported (valence) emotions. Nevertheless, service failures and experienced emotions need to be understood for various types of service encounters, including the more mundane ones. On the other hand, it is important to acknowledge that published academic research tends to be biased towards studies reporting significant relationships. By comparing both self-reported and neurophysiological metrics, the bigger (and real) picture can be made – resulting in a more powerful story with important implications for both theory and practice (cf. next section).

5.3. Theoretical and methodological contribution

This paper contributes to the affective CX and SST literature in several ways. The biggest contribution lies in the comparison of neurophysiological versus self-report measures to measure emotions and their dynamic nature during the CX. Many calls have been made to complement self-report data with neurophysiological data to study CX (Lemon and Verhoef, 2016; Venkatraman et al., 2012; Verhulst et al., 2019, 2020). Our study explored the merits of such an approach. When comparing the two different measures we see that for instance neurophysiological measures are better in picking up differences in arousal level across phases and did not overestimate valence during the different phases (e.g., Kim and Jang, 2014), unlike self-report phase-specific valence. Hence, the neurophysiological measures provide us with extra information on the CX.

Furthermore, customers often express a preference for either SST or HI, but it remains unclear if they affect emotions during the CX differently. Although a rich body of literature exists on SSTs and human service delivery in isolation, empirical studies comparing the different service interfaces are scarce (exception Köcher and Paluch, 2019). Our research fills this void by directly comparing these service interfaces by means of a real-life experiment. We find only a limited impact of human delivery compared with SST on the CX during a relatively mundane service encounter.

We should also look at the bigger picture, if we would have conducted a conventional study of emotions during the CX (i.e. not including neurophysiological measures) the conclusion of this study would be that service interfaces have an impact on emotions, since emotions (based on phase-specific valence) were less negative during service delivery by a human. However, we found no impact of human delivery versus SST in this context on the outcome variables. Thus, this study would probably never be published, since a null finding is often not considered interesting (cf. publication bias; van Witteloostuijn, 2016). Including the neurophysiological results, in fact, strengthens the outcome that service delivery by SST or an employee (at least in this

Table 4
Overview of results.

	NEUROPHYSIOLOGICAL		SELF-REPORT
Arousal			
Phase	Effect of Phase: goodbye > others; increased throughout the journey		Not applicable
Service interface	No effect of service interface		No effect of service interface
Valence			
Phase	No effect of Phase	≠	Effect of Phase: failure and service intervention 2 < other phases
Service interface	No effect of Service interface	≠	Effect of service interface^{+}: Human > SST Global valence no effect of service interface

Note. Results based on 2 (SST versus human, between subjects) × 6 (different phases, within subjects) ARH(1) mixed linear model analyses, (*) ANOVA used, ($^{+}$) marginally significant at 0.08.

context) does not matter. Emotions measured in-the-moment by EEG showed no difference in valence during specific phases and between human and SST encounters. This finding better reconciles with the found non-effect on the outcome variables (i.e., satisfaction and behavioral intentions). In conclusion, applying the neurophysiological measures added valuable insights to our understanding.

5.4. Managerial implications

We show that neurophysiological tools could be valuable for managers and service designers to help understand which moments are crucial during a CX because they allow visualizing how emotions vary during its different phases. In other words, these tools can help identify points where interventions might be necessary. However, keeping in mind the benefits, the associated extra effort and costs might not be worthwhile, especially for simple customer experiences. Nevertheless, we do suggest that for complex customer experiences that take longer (e.g., airport, festival), a neurophysiological approach could be valuable. During these experiences, phases may be less clear, and it might be harder to pinpoint important emotion-laden points in time with surveys.

As technology advances, an increasing number of companies are replacing frontline employees with SSTs. At the same time, creating valuable CXs is becoming increasingly important and insights into how to create valuable CXs in SST environments are lacking. Hence, our findings have several implications for managers. Our results suggest that real moments of truth, according to the customers' self-report data, are indeed during the failure phase. Furthermore, we show that in a relatively simple service encounter, even with a mild failure, a human employee could be easily replaced with an SST. Yet in more complex contexts, this approach might not work well. Indeed, in our study, both service interfaces were delivered in exactly the same way, whereas in real life, an employee can, for instance, behave in a way that builds rapport. This change in employee behavior could potentially make human service delivery better than an SST (e.g., Giebelhausen et al., 2014). In service contexts requiring high rapport (e.g., doctor's office, hairdresser; Gremler and Gwinner, 2000), we think that human employees can still make a difference.

5.5. Limitations and future research

This study sheds light on the impact of emotions (both valence and arousal) experienced by customers during a service encounter with a human employee versus SST, thereby taking the different measurement tools (EEG, GSR, and self-reported measures) into consideration. Some limitations, however, suggest directions for future research.

First, we showed that when neurophysiological measures are used, different results can surface compared with self-report measures. We have put forward a plausible explanation, specifically that positive and negative valence towards a service encounter can be overestimated by the use of self-report measures compared with other measures (e.g., Kim and Jang, 2014). Specifically, overestimation of emotions can be fused by a variety of processes linked to individual differences (e.g., awareness of own emotions; Salgado and Kingo, 2019), context (e.g., ordinary life versus intense event; Salgado and Kingo, 2019), other biases (e.g., halo bias (Wirtz, 2003), or even methodical artifacts (e.g., Robinson and Clore, 2002; Van Vaerenbergh and Thomas, 2013). Therefore, future research could focus on both why and when humans overestimate emotions or service performance outcomes would be a great way to stimulate debate on self-reported emotions and service performance outcomes in the service field.

Second, in some contexts, self-reported emotions might be accurate enough, whereas in other contexts complementary measures such as neurophysiological measures will be necessary (Harmon-Jones et al., 2016). It was beyond the scope of this paper, but future studies should compare self-report versus neurophysiological results focused on determining which type predicts future customer behavior and performance outcomes best and under which circumstances. Knowing when self-reports are good enough or even preferred is vital to investigate since they are cheaper, less intrusive, and easier to use compared to neuroscientific tools. Specifically, recent studies in other domains showed that neuroscientific tools hold potential to improve behavioral predictions on an individual and group level (e.g., Boksem and Smidts, 2014). For example, Genevsky et al. (2017) revealed that brain activation successfully predicted population-level and individual-level success of crowdfunding, while survey measures (e.g., affect ratings) could not. Furthermore, Boksem and Smidt (2014) showed that neuroscientific tools added unique variance in predicting movie trailer success, on top of self-reports. Thus, we advise future researchers to explore and combine both self-report and neuro-tools to advance prediction of service outcomes and other behavioral customer responses (cf., Bell et al., 2018; Lajante and Ladhari, 2019; Verhulst et al., 2019).

Third, in this paper, we started from the notion that emotions fall apart in two dimensions, specifically valence and arousal. Yet both on the theoretical and empirical level there have been discussions on including a third dimension, namely dominance (e.g., Mehrabian and Russell, 1974). Our rationale for focusing on the two-dimensional circumplex of emotion (e.g., Poels and Dewitte, 2019; Russell, 1980) was threefold and includes theoretical, methodological, and practical reasons. From a theoretical perspective, the dominance dimension has been considered to be more of a cognitive than an emotional response in past research (cf., Russell, 2003). From a methodological perspective, dominance often explains only a small part of the variance explained compared to arousal and valence (e.g., Bradley and Lang, 1994). Additionally, tools to measure both valence and arousal are well-validated, commonly used, and less complex (Jerram et al., 2014; Verhulst et al., 2019). From a practical perspective, and perhaps most important, the neuroscientific correlates of valence and arousal, unlike for dominance, are well understood (Jerram et al., 2014). Thus, most studies adopting neuroscientific tools rely on the two-dimensional approach. For an exception, see Jerram et al. (2014) who successfully adopts fMRI to study dominance. This being said, future research about the dominance dimension could definitely enrich our understanding of emotions during the CX and emotions in general.

Fourth, the self-reported method we used to assess pleasure and arousal (cf. The PAD scale) yielded poor psychometric results for arousal when including all six dichotomous items. Based on a confirmatory factor analysis, we had to delete three items of the original arousal items because of poor factor loadings (all lower than 0.5). Nevertheless, it is not uncommon for the PAD-scale to perform poorly, in particular on arousal. For example, Garaus and Wagner (2016) and Huang et al. (2017) faced the same issues. This may call for either the application of other pleasure and arousal scales, such as the Affect GRID (AIM; Russell et al., 1989), Profile of Moods Scale (POMS; McNair et al., 1971) or Positive and Negative Affect Schedule (PANAS; Watson et al., 1988). Or even the adoption and further development of new and innovative self-reported pleasure and arousal measures. For instance, scales that rely on emoji and emoticons seems very promising in that perspective (e.g., Lebender emoticon PANAVA' scale; Schreiber and Jenny, 2020).

Fifth, although using neuroscientific tools is common in domains such as psychology and organizational behavior, one could argue that participants are not used to wearing sensors on their forehead and shoulder. Consequently, this could lead to feeling a bit uncomfortable during the experiment and could decrease ecological validity. Previous research has suggested that wearing sensors on the face can be perceived as intrusive, yet it was only rated as moderately obstructive (Wrobel, 2018). To limit the impact of this uncomfortable feeling we first had participants watch a white cross on a screen and relax for 1 min 30 s before using the data for our baseline. Participants only started the service experience about 3–4 min after the sensors were installed, since some time passed people might be less prone to think about the sensors during the interaction. Therefore, another interesting approach would be to use mobile neurophysiological devices (e.g., Shimmer, 2020) or wearables (e.g., Fitbit; Maijers et al., 2018), during a natural (not experimental) CX to capture how customers feel along the customer journey (McColl-Kennedy et al., 2019; Verhulst et al., 2019). Specifically, using wearable and mobile tools might help to decrease potential awareness of wearing sensors throughout the CX.

Sixth, our study only focused on one single-touchpoint CX, so it does not allow generalizing to a CX that has several touchpoints (e.g., several visits at a store) or assessing long-term effects of service interfaces. In addition, a field study would increase the external validity of results. For example, we could investigate emotional changes during a CX linked to a real purchase decision journey. To illustrate, further research could, measure neurophysiological signals at the airport, starting from when a consumer arrives until boarding their flight. Such data could provide service researchers and practitioners with incredibly valuable information about the touchpoints and linked emotions that are crucial throughout the whole CX. In this regard, using wearable tools would allow boosting sample sizes of neuroscientific studies, since conducting studies with wearable sensors is less time-consuming and complex (e.g., Verhulst et al., 2019). Though it is common for this type of study (e.g., Boshoff, 2012), the sample size of our experiment is relatively small.

Seventh, our sample consisted mainly out of young adults. The literature suggests that emotional responses or experience change over peoples' lifespan (e.g., Carstensen et al., 2000). Future studies could, therefore, focus on emotions and attention for emotional stimuli during the CX in older samples. Comparing how emotions during a CX might fluctuate differently across age samples and how this is tied to behavioral outcomes or the overall service experience is a fruitful future research avenue (e.g., Alkire et al., 2020 this issue; Babin et al., 2020 this issue). Here neuroscientific tools might add an extra dimension, because some of those tools are particularly equipped to help understand individual differences and group differences better (Verhulst et al., 2019). Next to age, also culture could be integrated as a focal variable, to deepen our understanding of emotional responses during the CX (Rukavina et al., 2016).

Eighth, in this study we focused on the creation of a loyalty card and a computer that temporarily (15 s) could not save the data as the type of failure under investigation. In contrast with most prior literature, we did not find strong negative results of the failure. One plausible explanation could be that the failure while creating a loyalty card was not considered in a very negative light. Moreover, a computer that temporarily (15 s) could not save the data could be classified as a rather mild failure or not even considered as a failure at all (conform rather positive scenario valence scores). Some might even argue that creating a loyalty card is perhaps too simple for revealing large big emotional effects, and customers might not perceive it as a particularly important service. However, it is important to note that we investigated this type of service and failure because most service encounters are fairly positive or neutral and therefore deserve to receive attention as well. In addition, Schmitt et al. (2015) suggested that every service exchange leads to a CX, regardless of its nature and form. Therefore, we believe that the investigation of other types of failures, different levels of failure severity, and different types of recovery strategies offers another fruitful area for future research.

Ninth, in line with current trends, another interesting avenue would be to investigate how customers emotionally respond to blends of service interfaces where SST is combined with human features (e.g., anthropomorphism) or the impact of technology infusion such as (ro) bots (De Keyser et al., 2019; van Doorn et al., 2017; Wirtz et al., 2018).

Finally, mixing SST with frontline employees would also be interesting to investigate. Indeed, technology (and SST) can be seen on a continuum from totally replacing an employee to complementing an employee (Marinova et al., 2017). Investigating how SST and employees can work together to bring the CX to new heights is an interesting research opportunity.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jretconser.2020.102217.

References

Alkire, L., O'Connor, G.E., Myrden, S., Koecher, S., 2020. Patient experience in the digital age: an investigation into the effect of generational cohorts. J. Retail. Consum. Serv. This issue.

Anand, A., 2011. Welcome Valued Customer... to More Self-Checkouts. Nbcnews.com. http://www.nbcnews.com/id/43729757/ns/business-retail/t/welcome-valued-customer-more-self-checkouts/#.XAaHqWhKjcs. (Accessed 7 January 2020).

Appel, J., Von Der Pütten, A., Krämer, N.C., Gratch, J., 2012. Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. Adv. Hum. Comput. Interact. 324694. https://doi.org/10.1155/2012/324694.

Babin, B., Zhuang, W., Borges, A.J., 2020. Managing service recovery experience: effects of the forgiveness for older consumers. J. Retailing Consum. Serv. This issue.

Barrett, L.F., Mesquita, B., Ochsner, K.N., Gross, J.J., 2007. The experience of emotion. Annu. Rev. Psychol. 58, 373–403. https://doi.org/10.1146/annurev.psych.58.110405.085709.

Baumeister, R.F., Bratslavsky, E., Finkenauer, C., Vohs, K.D., 2001. Bad is stronger than good. Rev. Gen. Psychol. 5, 323–370. https://doi.org/10.1037//1089-2680.5.4.323.

Bell, L., Vogt, J., Willemse, C., Routledge, T., Butler, L.T., Sakaki, M., 2018. Beyond self-report: a review of physiological and neuroscientific methods to investigate consumer behavior. Front. Psychol. 9, e1655. https://doi.org/10.3389/fpsyg.2018.01655.

Benoit, S., Scherschel, K., Ates, Z., Nasr, L., Kandampully, J., 2017. Showcasing the diversity of service research: theories, methods, and success of service articles. J. Serv. Manag. 28, 810–836. https://doi.org/10.1108/JOSM-05-2017-0102.

Bianchin, M., Angrilli, A., 2012. Gender differences in emotional responses: a psychophysiological study. Physiol. Behav. 105, 925–932. https://doi.org/10.1016/j.physbeh.2011.10.031.

Bigné, J.E., Mattila, A.S., Andreu, L., 2008. The impact of experiential consumption cognitions and emotions on behavioral intentions. J. Serv. Market. 22, 303–315. https://doi.org/10.1108/08876040810881704.

Bitner, M.J., 1992. Servicescapes: the impact of physical surroundings on customers and employees. J. Market. 56, 57–71. https://doi.org/10.2307/1252042.

Bitner, M.J., Booms, B.H., Tetreault, M.S., 1990. The service encounter: diagnosing favorable and unfavorable incidents. J. Market. 54, 71–84. https://doi.org/10.2307/1252174.

Blut, M., Wang, C., Schoefer, K., 2016. Factors influencing the acceptance of self-service technologies. J. Serv. Res. 19, 396–416. https://doi.org/10.1177/1094670516662352.

Boksem, M.a.S., Smidts, A., 2014. Brain responses to movie-trailers predict individual preferences for movies and their population-wide commercial success. J. Market. Res. 52, 482–492. https://doi.org/10.1509/jmr.13.0572.

Boshoff, C., 2012. A neurophysiological assessment of consumers' emotional responses to service recovery behaviors: the impact of ethnic group and gender similarity. J. Serv. Res. 15, 401–413. https://doi.org/10.1177/1094670512453879.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatr. 25, 49–59. https://doi.org/10.1016/0005-7916(94)90063-9.

Bruun, A., Ahm, S., 2015. Mind the gap! Comparing retrospective and concurrent ratings of emotion in user experience evaluation. In: Abascal, J., et al. (Eds.), Proceedings of the 15th Conference on Human-Computer Interaction. INTERACT, Bamberg, pp. 237–254. https://doi.org/10.1007/978-3-319-22701-6_17.

Cacioppo, J.T., 2002. Social neuroscience: understanding the pieces fosters understanding the whole and vice versa. Am. Psychol. 57, 819–831. https://doi.org/10.1037//0003-066x.57.11.819.

Carstensen, L.L., Pasupathi, M., Mayr, U., Nesselroade, J.R., 2000. Emotional experience in everyday life across the adult life span. J. Pers. Soc. Psychol. 79, 644–655. https://doi.org/10.1037/0022-3514.79.4.644.

Caruelle, D., Gustafsson, A., Shams, P., Lervik-Olsen, L., 2019. The use of electrodermal activity (EDA) measurement to understand consumer emotions – a literature review and a call for action. J. Bus. Res. 104, 146–160. https://doi.org/10.1016/j.jbusres.2019.06.041.

Chebat, J.-C., Slusarczyk, W., 2005. How emotions mediate the effects of perceived justice on loyalty in service recovery situations: an empirical study. J. Bus. Res. 58, 664–673. https://doi.org/10.1016/j.jbusres.2003.09.005.

Christopoulos, G.I., Uy, M.A., Yap, W.J., 2019. The body and the brain: measuring skin conductance responses to understand the emotional experience. Organ. Res. Methods 22, 394–420. https://doi.org/10.1177/1094428116681073.

Clore, G.L., Wyer Jr., R.S., Dienes, B., Gasper, K., Gohm, C., Isbell, L., 2001. Affective feelings as feedback: some cognitive consequences. In: Martin, L.L., Clore, G.L. (Eds.), Theories of Mood and Cognition: A User's Guidebook. Lawrence Erlbaum Associates Publishers, Mahwah, pp. 27–62.

Coan, J.A., Allen, J.J.B., 2003. Frontal EEG asymmetry and the behavioral activation and inhibition systems. Psychophysiology 40, 106–114. https://doi.org/10.1111/1469-8986.00011.

Collier, J.E., Sherrell, D.L., 2010. Examining the influence of control and convenience in a self-service setting. J. Acad. Market. Sci. 38, 490–509. https://doi.org/10.1007/s11747-009-0179-4.
Dabholkar, P.A., 1996. Consumer evaluations of new technology-based self-service options: an investigation of alternative models of service quality. Int. J. Res. Market. 13, 29–51. https://doi.org/10.1016/0167-8116(95)00027-5.
Dabholkar, P.A., Bagozzi, R.P., 2002. An attitudinal model of technology-based self-service: moderating effects of consumer traits and situational factors. J. Acad. Market. Sci. 30, 184–201. https://doi.org/10.1177/0092070302303001.
Dabholkar, P.A., Michelle Bobbitt, L., Lee, E., 2003. Understanding consumer motivation and behavior related to self-scanning in retailing. Int. J. Serv. Ind. Manag. 14, 59–95. https://doi.org/10.1108/09564230310465994.
Davidson, R.J., Ekman, P., Saron, C.D., Senulis, J.A., Friesen, W.V., 1990. Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology. I. J. Pers. Soc. Psychol. 58, 330–341. https://doi.org/10.1037/0022-3514.58.2.330.
De Keyser, A., Köcher, S., Alkire (née Nasr), L., Verbeeck, C., Kandampully, J., 2019. Frontline service technology infusion: conceptual archetypes and future research directions. J. Serv. Manag. 30, 156–183. https://doi.org/10.1108/JOSM-03-2018-0082.
De Keyser, A., Lemon, K.N., Klaus, P., Keiningham, T.L., 2015. A framework for understanding and managing the customer experience. Mark. Sci. Inst. Work. Pap. Ser. 15–121.
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. J. Serv. Res. forthcoming.
DeWitt, T., Nguyen, D.T., Marshall, R., 2008. Exploring customer loyalty following service recovery. J. Serv. Res. 10, 269–281. https://doi.org/10.1177/1094670507310767.
Djelassi, S., Diallo, M.F., Zielke, S., 2018. How self-service technology experience evaluation affects waiting time and customer satisfaction? A moderated mediation model. Decis. Support Syst. 111, 38–47. https://doi.org/10.1016/j.dss.2018.04.004.
Edvardsson, B., Frow, P., Jaakkola, E., Keiningham, T.L., Koskela-Huotari, K., Mele, C., Tombs, A., 2018. Examining how context change foster service innovation. J. Serv. Manag. 29, 932–955. https://doi.org/10.1108/JOSM-04-2018-0112.
Espina, C., Pérez, M., 2014. The effect of self-scanning on consumers' expectations and satisfaction: an exploratory study in a retail service setting. In: Dheeraj, S., Borna, S. (Eds.), Proceedings of the 2007 Academy of Marketing Science Annual Conference. Springer, Cham, pp. 308–313.
Fan, A., Wu, L., Mattila, A.S., 2016. Does anthropomorphism influence customers' switching intentions in the self-service technology failure context? J. Serv. Market. 30, 713–723. https://doi.org/10.1108/JSM-07-2015-0225.
Furrer, O., Kerguignas, J., Delcourt, C., Gremler, D., 2020. Twenty-seven years of service research: a literature review and research agenda. J. Serv. Market. https://doi.org/10.1108/JSM-02-2019-0078.
Garaus, M., Wagner, U., 2016. Retail shopper confusion: conceptualization, scale development, and consequences. J. Bus. Res. 69, 3459–3467. https://doi.org/10.1016/j.jbusres.2016.01.040.
Gaur, S.S., Herjanto, H., Makkar, M., 2014. Review of emotions research in marketing, 2002–2013. J. Retailing Consum. Serv. 21, 917–923. https://doi.org/10.1016/j.jretconser.2014.08.009.
Genevsky, A., Yoon, C., Knutson, B., 2017. When brain beats behavior: neuroforecasting crowdfunding outcomes. J. Neurosci. 37, 8625–8863. https://doi.org/10.1523/JNEUROSCI.1633-16.2017.
Giebelhausen, M., Robinson, S.G., Sirianni, N.J., Brady, M.K., 2014. Touch versus tech: when technology functions as a barrier or a benefit to service encounters. J. Market. 78, 113–124. https://doi.org/10.1509/jm.13.0056.
Gijsenberg, M.J., Van Heerde, H.J., Verhoef, P.C., 2015. Losses loom longer than gains: modeling the impact of service crises on perceived service quality over time. J. Market. Res. 52, 642–656. https://doi.org/10.1509/jmr.14.0140.
Gremler, D.D., Gwinner, K.P., 2000. Customer-employee rapport in service relationships. J. Serv. Res. 3, 82–104. https://doi.org/10.1177/109467050031006.
Ha, H.-Y., Perks, H., 2005. Effects of consumer perceptions of brand experience on the web: brand familiarity, satisfaction and brand trust. J. Consum. Behav. 4, 438–452. https://doi.org/10.1002/cb.29.
Harmon-Jones, C., Bastian, B., Harmon-Jones, E., 2016. The discrete emotions questionnaire: a new tool for measuring state self-reported emotions. PloS One 11, e0159915. https://doi.org/10.1371/journal.pone.0159915.
Harris, J.M., Ciorciari, J., Gountas, J., 2018. Consumer neuroscience for marketing researchers. J. Consum. Behav. 17, 239–252. https://doi.org/10.1002/cb.1710.
Harris, K.E., Grewal, D., Mohr, L.A., Bernhardt, K.L., 2006. Consumer responses to service recovery strategies: the moderating role of online versus offline environment. J. Bus. Res. 59, 425–431. https://doi.org/10.1016/j.jbusres.2005.10.005.
Harrison-Walker, L.J., 2019. The effect of consumer emotions on outcome behaviors following service failure. J. Serv. Market. 33, 285–302. https://doi.org/10.1108/JSM-04-2018-0124.
Huang, M., Ali, R., Liao, J., 2017. The effect of user experience in online games on word of mouth: a pleasure-arousal-dominance (PAD) model perspective. Comput. Hum. Behav. 75, 329–338. https://doi.org/10.1016/j.chb.2017.05.015.
Hulland, J., Baumgartner, H., Smith, K.M., 2018. Marketing survey research best practices: evidence and recommendations from a review of JAMS articles. J. Acad. Market. Sci. 46, 92–108. https://doi.org/10.1007/s11747-017-0532-y.
Imhof, G., Klaus, P., 2019. The dawn of traditional CX metrics? Examining satisfaction, EXQ, and WAR. Int. J. Mark. Res. 32, 1–16. https://doi.org/10.1177/1470785319848955.
Jackson, D.C., Mueller, C.J., Dolski, I., Dalton, K.M., Nitschke, J.B., Urry, H.L., Rosenkranz, M.A., Ryff, C.D., Singer, B.H., Davidson, R.J., 2003. Now you feel it, now you don't: frontal brain electrical asymmetry and individual differences in emotion regulation. Psychol. Sci. 14, 612–617. https://doi.org/10.1046/j.0956-7976.2003.psci_1473.x.
Janda, S., Ybarra, A., 2005. Do product and consumer characteristics affect the relationship between online experience and customer satisfaction? J. Internet Commer. 4, 133–151. https://doi.org/10.1300/j179v04n04_09.
Jani, D., Han, H., 2015. Influence of environmental stimuli on hotel customer emotional loyalty response: testing the moderating effect of the big five personality factors. Int. J. Hospit. Manag. 44, 48–57. https://doi.org/10.1016/j.ijhm.2014.10.006.
Jerram, M., Lee, A., Negreira, A., Gansler, D., 2014. The neural correlates of the dominance dimension of emotion. Psychiatr. Res. Neuroimaging. 221, 135–141.
Kahneman, D., Krueger, A.B., 2006. Developments in the measurement of subjective well-being. J. Econ. Perspect. 20, 3–24. https://doi.org/10.1257/089533006776526030.
Kahneman, D., Wakker, P.P., Sarin, R., 1997. Back to Bentham? Explorations of experienced utility. Q. J. Econ. 112, 375–406. https://doi.org/10.1162/003355397555235.
Kattara, H.S., El-Said, O.a., 2014. Customers' preferences for new technology-based self-services versus human interaction services in hotels. Tourism Hospit. Res. 13, 67–82. https://doi.org/10.1177/1467358413519261.
Kenning, P., Plassmann, H., Ahlert, D., 2007. Applications of functional magnetic resonance imaging for market research. Qual. Mark. Res. Int. J. 10, 135–152. https://doi.org/10.1108/13522750710740817.
Kim, C., Kim, S., Im, S., Shin, C., 2003. The effect of attitude and perception on consumer complaint intentions. J. Consum. Market. 20, 352–371. https://doi.org/10.1108/07363760310483702.
Kim, J.-H., Jang, S. (Shawn), 2014. A scenario-based experiment and a field study: a comparative examination for service failure and recovery. Int. J. Hospit. Manag. 41, 125–132. https://doi.org/10.1016/j.ijhm.2014.05.004.
Kirk, Roger E., 2003. Experimental design. In: Weiner, I.B. (Ed.), Handbook of Psychology, vol. 2. John Wiley & Sons, inc., Hoboken, New Jersey.
Klaus, P., 2014. Measuring Customer Experience: How to Develop and Execute the Most Profitable Customer Experience Strategies. Palgrave Macmillan, UK.
Köcher, S., Paluch, S., 2019. "My bad": investigating service failure effects in self-service and full-service settings. J. Serv. Market. 33, 181–191. https://doi.org/10.1108/JSM-03-2018-0096.
Kuo, Y.-F., Wu, C.-M., 2012. Satisfaction and post-purchase intentions with service recovery of online shopping websites: perspectives on perceived justice and emotions. Int. J. Inf. Manag. 32, 127–138. https://doi.org/10.1016/j.ijinfomgt.2011.09.001.
Kuppelwieser, V.G., Klaus, P., 2019. Call for special issue "Exploring customer experience dynamics. J. Retailing Consum. Serv.
Ladhari, R., 2009. Service quality, emotional satisfaction, and behavioural intentions: a study in the hotel industry. Manag. Serv. Qual. Int. J. 19, 308–331. https://doi.org/10.1108/09604520910955320.
Lajante, M., Ladhari, R., 2019. The promise and perils of the peripheral psychophysiology of emotion in retailing and consumer services. J. Retailing Consum. Serv. 50, 305–313. https://doi.org/10.1016/j.jretconser.2018.07.005.
Larivière, B., Bowen, D., Andreassen, T.W., Kunz, W., Sirianni, N.J., Voss, C., Wünderlich, N.V., De Keyser, A., 2017. "Service Encounter 2.0": an investigation into the roles of technology, employees and customers. J. Bus. Res. 79, 238–246. https://doi.org/10.1016/j.jbusres.2017.03.008.
Lazarus, R.S., 1995. Vexing research problems inherent in cognitive-mediational theories of emotion-and some solutions. Psychol. Inq. 6, 183–196. https://doi.org/10.1207/s15327965pli0603_1.
Lemon, K.N., Verhoef, P.C., 2016. Understanding customer experience throughout the customer journey. J. Market. 80, 69–96. https://doi.org/10.1509/jm.15.0420.
Lerner, J.S., Keltner, D., 2000. Beyond valence: toward a model of emotion-specific influences on judgement and choice. Cognit. Emot. 14, 473–493. https://doi.org/10.1080/026999300402763.
Liapis, A., Katsanos, C., Sotiropoulos, D., Xenos, M., Karousos, N., 2015. Recognizing emotions in human computer interaction: studying stress using skin conductance. In: IFIP Conference on Human-Computer Interaction. Springer, Cham, pp. 255–262. https://doi.org/10.1007/978-3-319-22701-6_18.
Liljander, V., Strandvik, T., 1997. Emotions in service satisfaction. Int. J. Serv. Ind. Manag. 8, 148–169. https://doi.org/10.1108/09564239710166272.
Lin, J.-S.C., Hsieh, P.-L., 2007. The influence of technology readiness on satisfaction and behavioral intentions toward self-service technologies. Comput. Hum. Behav. 23, 1597–1615. https://doi.org/10.1016/j.chb.2005.07.006.
Liu, W., Sparks, B., Coghlan, A., 2016. Measuring customer experience in situ : the link between appraisals, emotions and overall assessments. Int. J. Hospit. Manag. 59, 42–49. https://doi.org/10.1016/j.ijhm.2016.09.003.
Maijers, M.C., Verschuren, O., Stolwijk-Swüste, J.M., van Koppenhagen, C.F., de Groot, S., Post, M.W.M., 2018. Is Fitbit Charge 2 a feasible instrument to monitor daily physical activity and handbike training in persons with spinal cord injury? A pilot study. Spinal Cord 4, 84. https://doi.org/10.1038/s41394-018-0113-4.
Manthiou, A., Hickmann, E., Klaus, P., 2020. Beyond good and bad: challenging the suggested role of emotions in customer experience (CX) research. J. Retailing Consum. Serv. This issue).
Marinova, D., de Ruyter, K., Huang, M.-H., Meuter, M.L., Challagalla, G., 2017. Getting Smart: learning from technology-empowered frontline interactions. J. Serv. Res. 20, 29–42. https://doi.org/10.1177/1094670516679273.
Mattila, A.S., Cho, W., Ro, H. (Cheyenne), 2011. The role of self-service technologies in restoring justice. J. Bus. Res. 64, 348–355. https://doi.org/10.1016/j.jbusres.2010.02.014.

Mauss, I.B., Robinson, M.D., 2009. Measures of emotion: a review. Cognit. Emot. 23, 209–237. https://doi.org/10.1080/02699930802204677.
Maxham, J.G., 2001. Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. J. Bus. Res. 54, 11–24. https://doi.org/10.1016/s0148-2963(00)00114-4.
McColl-Kennedy, J.R., Gustafsson, A., Jaakkola, E., Klaus, P., Radnor, Z.J., Perks, H., Friman, M., 2015. Fresh perspectives on customer experience. J. Serv. Market. 29, 430–435. https://doi.org/10.1108/JSM-01-2015-0054.
McColl-Kennedy, J.R., Zaki, M., Lemon, K.N., Urmetzer, F., Neely, A., 2019. Gaining customer experience insights that matter. J. Serv. Res. 22, 8–26. https://doi.org/10.1177/1094670518812182.
McNair, D., Lorr, M., Droppleman, L., 1971. Manual for the Profile of Mood States. Educational and Industrial Testing Service., San Diego.
Mehrabian, A., Russell, J.A., 1974. An Approach to Environmental Psychology. The MIT Press, Cambridge, MA.
Meuter, M.L., Ostrom, A.L., Roundtree, R.I., Bitner, M.J., 2000. Self-service technologies: understanding customer satisfaction with technology-based service encounters. J. Market. 64, 50–64. https://doi.org/10.1509/jmkg.64.3.50.18024.
Myrtek, M., Brügner, G., 1996. Perception of emotions in everyday life: studies with patients and normals. Biol. Psychol. 42, 147–164. https://doi.org/10.1016/0301-0511(95)05152-X.
Oh, H., Jeong, M., Baloglu, S., 2013. Tourists' adoption of self-service technologies at resort hotels. J. Bus. Res. 66, 692–699. https://doi.org/10.1016/j.jbusres.2011.09.005.
Ohme, R., Matukin, M., Pacula-Lesniak, B., 2011. Biometric measures for interactive advertising research. J. Interact. Advert. 11, 60–72. https://doi.org/10.1080/15252019.2011.10722185.
Ohme, R., Reykowska, D., Wiener, D., Choromanska, A., 2009. Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures. J. Neurosci. Psychol. Econ. 2, 21–31. https://doi.org/10.1037/a0015462.
Ostrom, A.L., Parasuraman, A., Bowen, D.E., Patricio, L., Voss, C.A., 2015. Service research priorities in a rapidly changing context. J. Serv. Res. 18, 127–159. https://doi.org/10.1177/1094670515576315.
Petzer, D.J., De Meyer, C.F., Svari, S., Svensson, G., 2012. Service receivers' negative emotions in airline and hospital service settings. J. Serv. Market. 26, 484–496. https://doi.org/10.1108/08876041211266413.
Pizzi, G., Marzocchi, G.L., Orsingher, C., Zammit, A., 2015. The temporal construal of customer satisfaction. J. Serv. Res. 18, 484–497. https://doi.org/10.1177/1094670515584752.
Pocock, S.J., Assmann, S.E., Enos, L.E., Kasten, L.E., 2002. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat. Med. 21, 2917–2930. https://doi.org/10.1002/sim.1296.
Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J. Appl. Psychol. 88, 879–903. https://doi.org/10.1037/0021-9010.88.5.879.
Podsakoff, P.M., Organ, D.W., 1986. Self-reports in organizational research: problems and prospects. J. Manag. 12, 531–544. https://doi.org/10.1177/014920638601200408.
Poels, K., Dewitte, S., 2019. The role of emotions in advertising: a call to action. J. Advert. 48, 81–90. https://doi.org/10.1080/00913367.2019.1579688.
Poels, K., Dewitte, S., 2006. How to capture the heart? Reviewing 20 years of emotion measurement in advertising. J. Advert. Res. 46, 18–37. https://doi.org/10.2501/S0021849906060041.
Puccinelli, N.M., Goodstein, R.C., Grewal, D., Price, R., Raghubir, P., Stewart, D., 2009. Customer experience management in retailing: understanding the buying process. J. Retailing 85, 15–30. https://doi.org/10.1016/j.jretai.2008.11.003.
Robertson, N., McDonald, H., Leckie, C., McQuilken, L., 2016. Examining customer evaluations across different self-service technologies. J. Serv. Market. 30 https://doi.org/10.1108/JSM-07-2014-0263.
Robinson, M.D., Clore, G.L., 2001. Simulation, scenarios, and emotional appraisal: testing the convergence of real and imagined reactions to emotional stimuli. Pers. Soc. Psychol. Bull. 27, 1520–1532. https://doi.org/10.1177/01461672012711012.
Robinson, M.D., Clore, G.L., 2002. Belief and feeling: evidence for an accessibility model of emotional self-report. Psychol. Bull. 28, 934–960. https://doi.org/10.1037/0033-2909.128.6.934.
Rukavina, S., Gruss, S., Hoffmann, H., Tan, J.-W., Walter, S., Traue, H.C., 2016. Affective computing and the impact of gender and age. PloS One 11, e0150584. https://doi.org/10.1371/journal.pone.0150584.
Russell, J.A., 2003. Core affect and the psychological construction of emotion. Psychol. Rev. 110, 145–172. https://doi.org/10.1037//0033-295x.110.1.145.
Russell, J.A., 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39, 1161–1178. https://doi.org/10.1037/h0077714.
Russell, J.A., Weiss, A., Mendelsohn, G.A., 1989. Affect grid: a single-item scale of pleasure and arousal. J. Pers. Soc. Psychol. 57, 493. https://doi.org/10.1037/0022-3514.57.3.493.
Salgado, S., Kingo, O.S., 2019. How is physiological arousal related to self-reported measures of emotional intensity and valence of events and their autobiographical memories? Conscious. Cogn. 75, 102811. https://doi.org/10.1016/j.concog.2019.102811.
Schmitt, B., Joško Brakus, J., Zarantonello, L., 2015. From experiential psychology to consumer experience. J. Consum. Psychol. 25, 166–171. https://doi.org/10.1016/j.jcps.2014.09.001.
Schreiber, M., Jenny, G.J., 2020. Development and validation of the 'Lebender emoticon PANAVA' scale (LE-PANAVA) for digitally measuring positive and negative activation, and valence via emoticons. Pers. Indiv. Differ. 160, 109923. https://doi.org/10.1016/j.paid.2020.109923.
Shankar, V., Smith, A., Rangaswamy, a, 2003. Customer satisfaction and loyalty in online and offline environments. Int. J. Res. Market. 20, 153–175. https://doi.org/10.1016/S0167-8116(03)00016-8.
Shao, C.Y., Baker, J.A., Wagner, J., 2004. The effects of appropriateness of service contact personnel dress on customer expectations of service quality and purchase intention. J. Bus. Res. 57, 1164–1176. https://doi.org/10.1016/s0148-2963(02)00326-0.
Shi, P., 2011. Research of software human-computer interface based on cognitive psychology. In: Jin, D., Lin, D. (Eds.), Advances in Multimedia, Software Engineering and Computing, vol. 2. Springer, Berlin, Heidelberg, pp. 611–616. https://doi.org/10.1007/978-3-642-25986-9_95.
Shimmer, 2020. available at: www.shimmersensing.com. (Accessed 28 May 2020).
Singh, J., Brady, M., Arnold, T., Brown, T., 2017. The emergent field of organizational frontlines. J. Serv. Res. 20, 3–11. https://doi.org/10.1177/1094670516681513.
Solomon, M.R., Surprenant, C., Czepiel, J.A., Gutman, E.G., 1985. A role theory perspective on dyadic interactions: the service encounter. J. Market. 49, 99–111. https://doi.org/10.2307/1251180.
Tivadar, R.I., Murray, M.M., 2019. A primer on electroencephalography and event-related potentials for organizational neuroscience. Organ. Res. Methods 22, 69–94. https://doi.org/10.1177/1094428118804657.
Tombs, A., McColl-Kennedy, J.R., 2003. Social-servicescape conceptual model - google scholar. Market. Theor. 3, 447–475. https://doi.org/10.1177/1470593103040785.
Tronvoll, B., 2011. Negative emotions and their effect on customer complaint behaviour. J. Serv. Manag. 22, 111–134. https://doi.org/10.1108/09564231111106947.
Turel, O., Yuan, Y., Connelly, C.E., 2008. In justice we trust: predicting user acceptance of e-customer services. J. Manag. Inf. Syst. 24, 123–151. https://doi.org/10.2753/mis0742-1222240405.
Vakulenko, Y., Shams, P., Hellström, D., Hjort, K., 2019. Service innovation in e-commerce last mile delivery: mapping the e-customer journey. J. Bus. Res. 101, 461–468. https://doi.org/10.1016/j.jbusres.2019.01.016.
van Doorn, J., Mende, M., Noble, S.M., Hulland, J., Ostrom, A.L., Grewal, D., Petersen, J. A., 2017. Domo Arigato Mr. Roboto: emergence of automated social presence in organizational frontlines and customers' service experiences. J. Serv. Res. 20, 43–58. https://doi.org/10.1177/1094670516679272.
Van Vaerenbergh, Y., Thomas, T.D., 2013. Response styles in survey research: a literature review of antecedents, consequences, and remedies. Int. J. Publ. Opin. Res. 25, 195–217. https://doi.org/10.1093/ijpor/eds021.
van Witteloostuijn, A., 2016. What happened to Popperian falsification? Publishing neutral and negative findings: moving away from biased publication practices. Cross Cult. Strat. Manag. 23, 481–508. https://doi.org/10.1108/CCSM-03-2016-0084.
Venkatesh, V., Thong, J.Y.L., Xu, X., 2012. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q. 36, 157. https://doi.org/10.2307/41410412.
Venkatraman, V., Clithero, J.A., Fitzsimons, G.J., Huettel, S.A., 2012. New scanner data for brand marketers: how neuroscience can help better understand differences in brand preferences. J. Consum. Psychol. 22, 143–153. https://doi.org/10.1016/j.jcps.2011.11.008.
Verhoef, P.C., Antonides, G., de Hoog, A.N., 2004. Service encounters as a sequence of events. J. Serv. Res. 7, 53–64. https://doi.org/10.1177/1094670504266137.
Verhoef, P.C., Lemon, K.N., Parasuraman, A., Roggeveen, A., Tsiros, M., Schlesinger, L. A., 2009. Customer experience creation: determinants, dynamics and management strategies. J. Retailing 85, 31–41. https://doi.org/10.1016/j.jretai.2008.11.001.
Verhulst, N., De Keyser, A., Gustafsson, A., Shams, P., Van Vaerenbergh, Y., 2019. Neuroscience in service research: an overview and discussion of its possibilities. J. Serv. Manag. 30, 621–649. https://doi.org/10.1108/JOSM-05-2019-0135.
Verhulst, N., Vermeir, I., Slabbinck, H., 2020. Viewpoint: using neuroscience tools in service. J. Serv. Market. Forthcoming.
Verleye, K., 2015. The co-creation experience from the customer perspective: its measurement and determinants. J. Serv. Manag. 26, 321–342. https://doi.org/10.1108/JOSM-09-2014-0254.
Wang, Z., Arndt, A.D., Singh, S.N., Biernat, M., Liu, F., 2013. "You lost me at hello": how and when accent-based biases are expressed and suppressed. Int. J. Res. Market. 30, 185–196. https://doi.org/10.1016/j.ijresmar.2012.09.004.
Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. 54, 1063. https://doi.org/10.1037//0022-3514.54.6.1063.
Weijters, B., Rangarajan, D., Falk, T., Schillewaert, N., 2007. Determinants and outcomes of customers' use of self-service technology in a retail setting. J. Serv. Res. 10, 3–21. https://doi.org/10.1177/1094670507302990.
Wirtz, J., 2003. Halo in customer satisfaction measures : the role of purpose of rating, number of attributes and customer involvement. Int. J. Serv. Ind. Manag. 14, 96–119. https://doi.org/10.1108/09564230310466001.
Wirtz, J., Patterson, P.G., Kunz, W.H., Gruber, T., Lu, V.N., Paluch, S., Martins, A., 2018. Brave new world: service robots in the frontline. J. Serv. Manag. 29, 907–931. https://doi.org/10.1108/JOSM-04-2018-0119.
Wrobel, M.R., 2018. Applicability of emotion recognition and induction methods to study the behavior of programmers. Appl. Sci. 8, 323. https://doi.org/10.3390/app8030323.

Yan, A., Solomon, S., Mirchandani, D., Lacity, M., Porra, J., 2013. The role of service agent, service quality, and user satisfaction in self-service technology. In: Thirty Fourth Int. Conf. Inf. Syst. Milan 2013, pp. 1–21.

Zhang, S., Paul, J., Nantha-Aree, M., Buckley, N., Shahzad, U., Cheng, J., DeBeer, J., Winemaker, M., Wismer, D., Punthakee, D., Avram, V., Thabane, L., 2014. Empirical comparison of four baseline covariate adjustment methods in analysis of continuous outcomes in randomized controlled trials. Clin. Epidemiol. 6, 227–235. https://doi.org/10.2147/CLEP.S56554.

Zomerdijk, L.G., Voss, C.A., 2010. Service design for experience-centric services. J. Serv. Res. 13, 67–82. https://doi.org/10.1177/1094670509351960.

Further reading

Bagozzi, P., Yi, Y., 1988. On the evaluation of structural equation models. J. Acad. Mark. 16, 74–94. https://doi.org/10.1007/BF02723327.

Ciorciari, J., 2012. Bioelectrical signals: the electroencephalogram (2012). In: Wood, A. (Ed.), Physiology, Biophysics and Biomedical Engineering. Taylor & Francis Inc, Washington, pp. 539–267.

Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. J. Market. Res. 18, 39–50. https://doi.org/10.2307/3151312.

Hu, L.T., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equ. Model. 6, 1–55. https://doi.org/10.1080/10705519909540118.

Netemeyer, R.G., Bearden, W.O., Sharma, S., 2003. Scaling Procedures: Issues and Applications. Sage Publications, Thousand Oaks. https://doi.org/10.4135/9781412985772.

Usakli, A.B., 2010. Improvement of EEG signal acquisition: an electrical aspect for state of the Art of front end. Comput. Intell. Neurosci. 2010 https://doi.org/10.1155/2010/630649.

van Dooren, M., de Vries, J.J.G., Janssen, J.H., 2012. Emotional sweating across the body: comparing 16 different skin conductance measurement locations. Physiol. Behav. 106, 298–304. https://doi.org/10.1016/j.physbeh.2012.01.020.

Nanouk Verhulst is working as a doctor assistant in the Business department at the Vrije Universiteit Brussels (VUB). She obtained a Ph.D. in Applied Business Economics (Ghent University). Further, she holds a Master's in organizational and occupational psychology (Ghent University) and a Master's in Management (KU Leuven). Her publications appeared in the Journal of Services Marketing and the Journal of Service Management.

Iris Vermeir obtained a master degree in Psychology in 1998 and a PhD in Applied economics in 2003 and works as associate professor at Ghent University. Her research focusses on several aspects of consumer behavior and marketing communication like attitude-behavior consistency, sustainable consumption, (ethical) decision making, and (healthy) food consumption. She has presented her research on several international conferences and published in various top-tier scientific journals in the field of Marketing, Communication, Ethics, Psychology, Agricultural Sciences, and Food Science. Journals include amongst others Journal of Consumer Research, Journal of Consumer Psychology, International Journal of Advertising, Journal of Business Research, Psychology & Marketing, Journal of Economic Psychology, Ecological Economics, Journal of Agricultural and Environmental Ethics, Appetite and Food Quality and Preference. She is member of the board of the journal Sustainability and ad hoc reviewer for several international scientific journals like Appetite, Journal of Business Ethics, Journal of Economic Psychology, Journal of Consumer Behavior, Foods, and Ecological Economics. She is also head of the research center BE4lLIFE 'behavioral economics for life', an interdisciplinary Ugent research center that works together on increasing well-being enhancing consumption (i.e. sustainable, healthy, and ethical consumption).

Hendrik Slabbinck is Associate Professor at Ghent University in Belgium. After obtaining his PhD in Applied Economics in 2012, he was appointed Assistant Professor at the University College Ghent. He joined the Department of Marketing, Innovation, and Organization at Ghent University in October 2013 and was appointed as Associate Professor in October 2018. Hendrik is a research fellow at the Be4Life research center that stimulates and integrates academic and applied behavioral research on topics that have a positive impact on the well-being of society. His research centers around the development and application of implicit measures in management and business research. He is ad hoc reviewer for many journals such as Small Business Economics, Journal of Consumer Behavior, and PLOS One. His research has been published in PLOS One, Journal of Business Research, Journal of Experimental and Social Psychology, Business Strategy and the Environment, Environment and Behavior, among others.

Bart Larivière is Associate Professor of Marketing at KU Leuven (Belgium) and Research Fellow at the Center for Service Intelligence (Ghent University). He holds a PhD degree from Ghent University. Bart's research focuses on the role of customers, employees, and technology in service encounters. Bart is passionate about both service marketing and digital marketing. Bart is also the current co-chair of the Service Research Special Interest Group (SERVSIG) of the American Marketing Association (AMA) and Research Editor for the Journal of Service Management.

Maurizio Mauri holds a Degree in Psychology, holds PhD from IULM University, and a post-doc at Università RomaTre in Neuromanagement. With several publications and as invited speaker in different international conferences, he is now Head of User Experience & Neuromarketing Research at SR LABS (eye-tracking company in Milan), and Adjunct Professor at Università Cattolica of Milan with a teaching position for the class "Cognitive Psychology applied to Ergonomics, User & Customer Experience. He also received a fellowship from the MIT-Italy Program to visit ('05-'07) the "Brain and Cognitive Sciences Dpt." at MIT to study how neuro and psychophysiological measures enable to assess emotional states. During his post-doc at IULM University ('08-'12), he works on a scientific model enabling to characterize quantitatively the level of emotional engagement. In '13-'17, he was Operational Coordinator at the "Behavior and Brain Lab" at IULM University.

Vincenzo Russo is Associate Professor of Consumer Psychology and Neuromarketing at IULM University in Milan. Director of the Center of Research for Neuromarketing, "Behavior and Brain Lab – IULM University. From 2012 He is Director of the Master in Food and Wine Communication and from 2009 He is Director of the Master in Sanitary Professions Coordinators and from 2016 Director of the Master in Sport Communication and Sport Marketing. Author of several papers and books about the Neuromarketing, Social Communication, and about the change management in the Not for Profit Organization. He has widely published in recognized international journals including the Journal of Consumer Behavior, Journal of Advertising Research, Food Quality and Preference; Journal of Global Information Management, European Journal of Information Systems, in the Italian journal of organizational psychology Risorsa Uomo.