Emotional arousal in customer experience: A dynamic view

Delphine Caruelle a, b, Poja Shams b, Anders Gustafsson c, Line Lervik-Olsen c

a Department of Marketing, Kristiania University College, Kirkegata 24-26, 0153 Oslo, Norway
b Service Research Center (CTF), Karlstad University, Universitetsgatan 2, 651 88 Karlstad, Sweden
c Department of Marketing, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway

ARTICLE INFO
Keywords:
Arousal
Customer experience
Emotion
Service encounters

ABSTRACT
Customer emotion in services has been extensively studied, but prior research has overlooked the dynamics of emotion over time. Our research addresses this gap by studying how emotional arousal varies throughout a service encounter. Drawing from the psychology literature, we identify certain features (or patterns) that characterize how arousal varies throughout a service encounter and predict how they may affect customer approach response (e.g., spending, unplanned purchases). We explore the effect of these features in field studies in two stores using a psychophysiological measure (electrodermal activity) to capture arousal over time. We find that (1) the highest arousal level reached during the encounter and (2) the skewness of the distribution of arousal levels (i.e., the frequency of lower arousal levels relative to higher ones) predict customer approach response. This paper opens new avenues for understanding customers from an emotional perspective, which can improve the customer experience in service encounters.

1. Introduction
Improving customer experience is a priority for many firms (Zendesk, 2023), and two-thirds of marketing managers recognize that their firms compete on customer experience (Pemberton, 2018). Firms no longer gain a competitive advantage solely through product quality or functionality; better than the emotional experience, as they lack the knowledge and tools to explore the emotional aspect (De Chernatony, 2010).

Service encounters give firms opportunities to emotionally engage with customers. During a service encounter, which essentially can be seen as a process during which customers interact with the firm’s offering, customers are exposed to multiple stimuli that elicit emotional responses and lead to various behaviors (Bagozzi et al., 1999; Bitner, 1992; Tombs & McColl-Kennedy, 2003). It is thus not surprising that the role of customer emotion in service encounters has been extensively studied (e.g., Chebat & Slusarczyk, 2005; Liljander & Mattsson, 2002; Liu et al., 2019; Mattila & Enz, 2002; Price et al., 1995; Snyder et al., 2022; Sun et al., 2023; Van Dolen et al., 2004).

However, prior research has paid little attention to the fact that emotion varies over time in a service encounter. With a few exceptions (see Table 1), marketing research has adopted a static view of customer emotion in service encounters. In most cases, the emotional state is measured at a single point in time, typically after the service encounter. Thus, little is known about the impact of variations in emotional state throughout a service encounter.

In this research, our primary objective is to examine the fluctuations in customers’ arousal levels throughout a service encounter (with arousal corresponding to the intensity of the emotional state; Lang et al., 1993; Russell, 1980) and their impact on customers’ responses to the service environment (i.e., servicescape; Bitner, 1992). Drawing on the fundamental principles established by Mehrabian and Russell (1974), we acknowledge that the multitude, scope, and interplay of stimuli in a servicescape can significantly affect customers’ emotional states and
Table 1: Summary of studies embracing a dynamic view of customer emotion in service encounters.

<table>
<thead>
<tr>
<th>Study</th>
<th>Data</th>
<th>Emotion measurement method</th>
<th>Theoretical model</th>
<th>Key findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>The present research</td>
<td>Field study in two stores, N = 125</td>
<td>Physiological (EDA)</td>
<td>✓ ✓ ✓ ✓</td>
<td>The peak (most emotionally arousing moment of the service encounter) increases customer approach response. Conversely, a positively skewed distribution of arousal levels during the encounter decreases customer approach response.</td>
</tr>
<tr>
<td>Boshoff (2012)</td>
<td>Lab scenario-based experiment, N = 64</td>
<td>Physiological (EDA for arousal, EEG and EMG for valence)</td>
<td>✓ ✓ ✓ ✓</td>
<td>Arousal does not significantly vary throughout the different phases of a service encounter comprising a service failure and recovery, whereas valence does.</td>
</tr>
<tr>
<td>Delakas (2006)</td>
<td>Lab scenario-based experiment, N = 127</td>
<td>Affect rating dial</td>
<td>✓ ✓ ✓ ✓</td>
<td>An improving trend in instances of emotions (first negative, then positive) throughout a service encounter yields higher satisfaction with the encounter than a deteriorating trend (first positive, then negative). In addition, experiencing a positive (vs. negative) emotion at the end of the encounter increases satisfaction.</td>
</tr>
<tr>
<td>Dubé and Morgan (1996)</td>
<td>Field study in a hospital, N = 96</td>
<td>Survey</td>
<td>✓ ✓ ✓ ✓</td>
<td>The emotional trend throughout the hospital stay does not significantly predict satisfaction with the stay.</td>
</tr>
<tr>
<td>Study</td>
<td>Data</td>
<td>Emotion measurement method</td>
<td>Theoretical model</td>
<td>Key findings</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------</td>
<td>-----------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Dubé and Morgan (1998)</td>
<td>Field study in a hospital, N = 93</td>
<td>Survey</td>
<td>✓ ✓ ✓ ✓</td>
<td>Gender and perceived health status predict the change in emotions over time during a hospital stay.</td>
</tr>
<tr>
<td>Maguire and Geiger (2015)</td>
<td>Field study across 13 service industries, N = 57</td>
<td>SMS diary</td>
<td>✓ ✓ ✓ ✓</td>
<td>Different emotions are experienced at different points in time during a service encounter.</td>
</tr>
<tr>
<td>McColl-Kennedy et al. (2019)</td>
<td>Longitudinal survey in a B2B context, N = 3116</td>
<td>Survey</td>
<td>✓ ✓ ✓ ✓</td>
<td>Customers experience various discrete emotions at touchpoints with the firm; joy and love are the two most frequent ones.</td>
</tr>
<tr>
<td>Verhoef et al. (2004)</td>
<td>Secondary data (phone calls with customer service), N = 97</td>
<td>Language analysis</td>
<td>✓ ✓ ✓ ✓</td>
<td>Peak emotion positively affects the evaluation of a service encounter, whereas emotion at the end of the encounter has a negative effect. Furthermore, an increasing trend in emotion yields a higher overall positive evaluation than a decreasing trend.</td>
</tr>
<tr>
<td>Verhulst et al. (2020)</td>
<td>Lab experiment, N = 40</td>
<td>Physiological (EDA for arousal, EEG for valence) and survey</td>
<td>✓ ✓ ✓ ✓</td>
<td>Arousal (measured physiologically) varies throughout the stages of a service encounter comprising a service failure and recovery. Valence varies throughout the different stages of an encounter when measured with self-reports but not when measured physiologically (EEG).</td>
</tr>
</tbody>
</table>

As customers navigate diverse combinations of stimuli, the intensity of their emotional response, or arousal level, may vary significantly. Our aim is to uncover how those variations influence customer response to the servicescape. This exploration provides a focused lens on the complex interplay between the service environment, customer emotion, and customers’ attitudes and behaviors in that environment.

Drawing from the model of evaluation by moments (Fredrickson, 2000; Kahneman, 2000; Kahneman & Frederick, 2002) and range-frequency theory (Parducci, 1965, 1968, 1995), we explore the impact on customer response of four features characterizing how arousal varies over the course of a service encounter: (1) the peak (i.e., highest arousal level during the encounter), (2) the end (i.e., arousal level in the last part of a service encounter), (3) the trend (i.e., increase or decrease in arousal level over the course of a service encounter), and (4) the skewness of the distribution of arousal levels (i.e., frequency of lower arousal levels relative to higher ones throughout the service encounter). Building on Mehrabian and Russell (1974), we focus on customer approach response, which describes the tendency to explore, have a favorable attitude toward, and return to the service environment.

We test our predictions in a field study in two stores where participants’ arousal levels are measured in real time throughout the encounter using electrodermal activity (EDA), a psychophysiological method. We find that a greater peak increases customer approach response and that a more positively skewed distribution of arousal levels (with lower arousal levels being more frequent than relatively higher ones) decreases customer approach behavior, but our predictions regarding the end and trend in arousal level are not empirically supported.

Our research makes several significant contributions. First, it contributes to the customer experience literature (Becker & Jaakkola, 2020; De Keysers et al., 2020; Jaakkola et al., 2022; Pucinelli et al., 2009; Roggeveen et al., 2020; Spence et al., 2014) by advancing knowledge on the role of arousal (as a dimension of emotion) at touchpoints. Prior
research shows that emotion resulting from a service encounter touchpoint influences customer response to the servicescape. Our research adds to this body of knowledge by showing how fluctuations in arousal level over the course of a service encounter predict customers’ responses. Second, our research makes a methodological contribution by measuring arousal level with a psychophysiological method (EDA) in a field study, thereby demonstrating how to capture arousal in situ and over time. By doing so, we are able to capture in real-time the response to the actual stimuli experienced holistically at a service touchpoint. Last, our research suggests that service managers should orchestrate customer experiences so that moments of low arousal are less frequent than moments of relatively higher arousal, thereby avoiding a positively skewed distribution of arousal levels during the service touchpoint. Managers should also aim at maximizing the highest arousal level reached at some point in the service encounter.

2. Literature review

2.1. Arousal

Arousal, together with valence, is an underlying dimension of emotion (Mehrabian & Russell, 1974; Posner et al., 2005). It corresponds to the intensity of the emotional state, whereas valence describes how positive or negative the emotional state is. Dimensional theories of emotion (Russell, 1980, 2003; Russell & Barrett, 1999; Watson & Tellegen, 1985) posit that an emotional state is the combination of a certain degree of arousal and valence. For example, fear and sadness are both negative states, but they differ in arousal level, fear being more intense than sadness (Posner et al., 2005; Russell, 1980; Russell & Barrett, 1999).

Emotional states manifest through expression (telling or reporting how one feels), behavior (facial expressions, body language), and physiological changes (Bradley & Lang, 2000; Gross, 1998; Kring & Gordon, 1998; Lang, 1993; Mauss & Robinson, 2009). Thus, the arousal dimension of emotion is shown in what individuals indicate they feel (e.g., on the Affect Grid; Russell et al., 1989), their voice pitch (Mauss & Robinson, 2009), and physiological changes, such as changes in EDA (Lang et al., 1993).

As a fundamental component of emotion (Bagozzi et al., 1999), arousal has been extensively studied. The consumer behavior literature examines its role in product evaluation (Bettiga et al., 2017; Hagtvedt, 2015; Noseworthy et al., 2014) and ad evaluation (Gorn et al., 2001; Mano, 1997; Sanbonmatsu & Kardes, 1988; Steenkamp et al., 1996). Research on online consumer behavior reveals that arousal is expressed in online complaints (Herhausen et al., 2023) and consumer reviews (Yin et al., 2017) and transmitted to consumers who read them (Fox et al., 2018). The service marketing literature has studied arousal as a response to environmental stimuli and an antecedent to customer attitudes and behaviors in physical service environments (De Nisco & Warmaby, 2014; Donovan & Rosset, 1982; Donovan et al., 1994; Dubé et al., 1995; Kaltcheva & Weitz, 2006; Lunardo & Roux, 2015; Ridgway et al., 1990; Sherman et al., 1997; Sweeney & Wyer, 2002; Walsh et al., 2011; Wirtz & Bateson, 1999; Wirtz et al., 2000; Yalch & Spangenberg, 2000) and online ones (Ha & Lennon, 2010; Menon & Kahn, 2002). Other service marketing studies (Boshoff, 2012; Verhulst et al., 2020) examine the impact on arousal of specific events in the customer journey, such as service failure and service recovery.

2.2. Emotion during service encounters

To understand the role of emotion in service encounters, we draw on Mehrabian and Russell (1974), who posit that stimuli in a physical environment (e.g., a servicescape) influence peoples’ emotional states and consequently their attitudinal and behavioral responses to the environment. Mehrabian and Russell describe people’s responses on a continuum from approach to avoidance, defined “in a broad sense to include physical movement toward, or away from, an environment or stimulus, degree of attention, exploration, favorable attitudes such as verbally or nonverbally expressed preference or liking, approach to a task [...] , and approach to another person [...]” (p. 96). An approach response is characterized by a desire to stay in, explore, affiliate with, and return to the environment, whereas an avoidance response is the opposite: wishing to avoid and leave the environment.

Mehrabian and Russell propose that the two dimensions of emotion, valence and arousal, play a fundamental role in predicting the approach-avoidance response, contending that the positive valence of an emotional state is positively related to approach responses and that arousal augments this relationship. In a pleasant environment (one triggering positive rather than negative emotion), increased arousal increases the approach response. These predictions have been tested in service and retail research, with multiple studies investigating how customers’ emotional states (valence and arousal) in a servicescape impact their approach response, operationalized as spending, unplanned purchases, spending (more) time in the servicescape, and satisfaction (for a review, see Vieira, 2013).

The main limitation of this model is that it sees arousal as one global emotional state resulting from exposure to multiple stimuli, whereas the customer experience literature (Becker & Jaakkola, 2020; De Keyser et al., 2020; Jaakkola et al., 2022; Puccinelli et al., 2009; Roggeveen et al., 2020; Voorhees et al., 2017) has brought to light that service encounters are composed of multiple stimuli and that the interplay of these stimuli affects how customers respond. Because of the successive additions and/or withdrawals of stimuli over time during a service encounter, the combination of stimuli that are perceived or sensed by customers evolves dynamically. The evolving configuration of stimuli makes customer experience dynamic, making the customer response vary over time as well (Becker & Jaakkola, 2020). Thus, arousal level likely varies throughout the service encounter.

Thus, studying arousal as a global emotional state for the entire encounter may not correctly account for customers’ emotional experiences. Instead, features or patterns in how arousal varies over the course of a service encounter may better reflect how customers experienced the encounter. Two theories help explain how variations in arousal may give a better account for the emotional experience than one global assessment: the model of evaluation by moments (Fredrickson, 2000; Kahneman, 2000; Kahneman & Frederick, 2002) and range-frequency theory (Parducci, 1965, 1968, 1995).

2.3. Model of evaluation by moments

The model of evaluation by moments was developed after Fredrickson and Kahneman (1993) and Kahneman et al. (1993) found evidence that the sum of emotional responses experienced in an episode (a discrete experience, such as a service encounter) does not influence how individuals evaluate the entire episode. Thus, a greater accumulation of pleasant moments does not necessarily yield a more favorable evaluation of the episode that these moments constitute. Rather, certain features or patterns characterizing how the emotional state varies over time determine how the entire episode is evaluated.

The model of evaluation by moments posits that two moments of an episode predict how the entire episode is evaluated: the peak and the end (the so-called peak-end rule). The peak is the moment of the most intense emotional state of the entire episode (i.e., the moment of highest arousal during a service encounter). It is not a set moment; rather, what characterizes the peak is that no other moment during the entire episode is as emotionally intense as the peak is. The end—the other influential moment—refers to the last part of the episode and its emotional intensity (i.e., arousal level in the last part of a service encounter).

Kahneman (2000) explains the peak-end rule by advancing that the peak and end are perceived as representative snapshots of the whole episode. Elaborating further on this idea, Fredrickson (2000) proposes that the peak and end are particularly meaningful moments: the peak
reveals how good the experience can be and the end typically coincides with the achievement (or lack of it) of a goal that one seeks to achieve in the given episode.

The model of evaluation by moments also contends that the extent to which the emotional state increases or decreases in intensity throughout an episode (i.e., increase or decrease in arousal level throughout a service encounter) impacts the overall evaluation of the episode (Fredrickson, 2000; Kahneman, 2000). This feature is called the trend. Individuals more favorably evaluate episodes that improve (e.g., an increase in arousal during a positive episode) as opposed to deteriorate over time.

These features advanced by the model of evaluation by moments have been tested in prior research on service encounters, albeit with mixed findings. Dube and Morgan (1996) found no significant effect of the trend in emotional intensity on satisfaction with a hospital stay. Hansen and Danaher (1999) found that an improving (vs. deteriorating) trend in service performance positively affected the overall evaluation of the service encounter. Verhoef et al. (2004) tested the influence of the peak, end, and trend on evaluating a call to a call center; a higher peak and improving trend yielded a greater overall positive evaluation, whereas a more pleasant end of the call had the opposite effect. Dalakas (2006) found that people rated a restaurant experience as more satisfying when it ended with a positive rather than negative incident and when the episode improved rather than deteriorated over time.

2.4. Range-Frequency theory

The second theory that helps explain how variations in emotional state may better account for emotional experience is range-frequency theory (Parducci, 1965, 1968, 1995). At a general level, this theory contends that when a set of stimuli is experienced, each stimulus is judged relative to the other stimuli in the set. Specifically, each stimulus is judged relative to the two most extreme (highest and lowest) stimuli in the set (the range principle). For example, a $6 product is perceived as cheaper when part of a set of products ranging from $5 to $9 rather than from $3 to $7. Additionally, each stimulus is judged based on the frequency of stimuli at the lower and upper ends of the range (the frequency principle). For example, assuming a constant range of prices (e.g., $5–$7), a $6 product is perceived as cheaper when $7 products (the upper end of the range) are more frequent than $5 products (the lower end of the range).

Importantly, the corollary of the range and frequency principles is that the skewness of the distribution of the stimuli (i.e., the extent to which those at the upper end of the range are more or less frequent than those at the lower end) influences how the set of stimuli, as a whole, is evaluated (Smith et al., 1989; Tripp & Brown, 2016). This has important implications for episodes that trigger a series of emotional responses (Hsee & Tsai, 2008; Parducci, 1968). The overall evaluation of such an episode is influenced by the skewness of the distribution of emotional responses in this episode. When the emotional responses are positive, a negative skewness of the distribution of the emotional responses leads to a more favorable evaluation than a positive skewness. This is because, under negative skewness, moments with a high emotional intensity (i.e., arousal levels at the upper end of the range) are more frequent than moments with a lower intensity (i.e., arousal levels at the lower end of the range), giving an overall impression that most of the episode was emotionally engaging. Conversely, under positive skewness, moments with a high emotional intensity are infrequent relative to those with a lower intensity, making it salient that most of the episode is not very emotionally engaging. Thus, positive skewness (i.e., lower arousal levels being more frequent than relatively higher ones) deteriorates the evaluation of the episode.

For example, suppose that visitors of an amusement park rated their experience of each attraction from 1 to 10 as a function of how emotionally arousing the attraction felt (with a higher rating indicating higher arousal). Person A and Person B experienced four attractions each; Person A rates them 4, 4, 4, and 10, and Person B rates them 1, 7, 7, and 7. The average rating for both is 5.5, but the skewness of the distribution of these ratings differs. The skewness for Person A is positive (equal to 2), meaning the arousing attraction (rated 10) was outnumbered (and thus overshadowed) by the less arousing ones (rated 4). Consequently, Person A’s overall impression is that most of the visit was not very exciting. By contrast, the skewness for Person B is negative (–2), meaning the attraction that was little arousing (rated 1) was outnumbered by the more arousing ones (rated 7). The single attraction rated 1 made it salient that the other attractions were highly arousing. Range-frequency theory predicts that Person B will more favorably evaluate the visit than Person A, although the average rating of the attractions they experienced is identical.

2.5. Hypotheses

Drawing from these theories, we propose that arousal and its variations during a service encounter influence customer approach response. The underlying premise of both the model of evaluation by moments and range-frequency theory is that variations in what individuals experience throughout an episode impact how they respond to that episode. Given this premise, we propose to apply the predictions made by the model of evaluation by moments and by range-frequency theory to customer approach response.

Specifically, building on the predictions made by the model of evaluation by moments (Fredrickson, 2000; Kahneman, 2000; Kahneman & Frederick, 2002), we propose that three features characterizing how arousal varies throughout a service encounter predict customer approach response: the peak (the highest arousal level reached at some point during the service encounter), the end (the arousal level during the last part of the service encounter), and the trend (the change in arousal level over the course of the service encounter). An approach response, as opposed to avoidance, implies that an encounter is pleasant, thus generating positive, rather than negative, emotion; so, assuming the context of a pleasant service encounter, we hypothesize the following:

H1: A greater peak arousal level has a positive effect on customer approach response.
H2: A greater end arousal level has a positive effect on customer approach response.
H3: An increasing trend in arousal level has a positive effect on customer approach response.

Range-frequency theory suggests that the skewness of the distribution of arousal levels influences approach response, so we predict that, for a pleasant service encounter, a positively skewed distribution of arousal levels decreases customer approach response. More formally, and assuming the context of a pleasant service encounter, we hypothesize:

H4: A more positively skewed distribution of arousal levels (i.e., lower arousal levels being more frequent than relatively higher arousal levels) has a negative effect on customer approach response.

3. Method

We tested our hypotheses in a field study, a strategic choice to retain the multifaceted nature of real-world servicescape stimuli, such as colors, smells, sounds, crowds, and sales promotions. These stimuli are integral to a holistic customer experience, and recreating such a rich, dynamic environment in a laboratory setting would pose great challenges. Aligning with Lemon and Verhoef’s emphasis (2016, p. 88) on in situ data collection in customer experience research, our methodology ensured a realistic, comprehensive examination of the factors under study.

The study was conducted in two stores (a hypermarket and a...
hardware store). Both are open behavior settings (Foxall & Greenley, 1999), meaning that customers have some control over the unfolding of their service encounters. This was particularly suitable for our research purpose, as approach response necessitates that customers have discretion over their actions in the service environment.

Data were first collected in a hypermarket in Sweden on December 13–20, 2016 (except on December 14). Research assistants recruited 101 persons at the store’s entrance, offering two lottery tickets for their participation. Data were next collected in a hardware store, also located in Sweden, on February 20–27, 2017 (except on February 26). Sixty participants were recruited in the same manner and with the same incentive as before.

3.1. Arousal measurement

We chose EDA—also known as skin conductance or galvanic skin response—as a physiological measure of arousal for two reasons. First, it is a highly valid, reliable metric for emotional arousal as evidenced by numerous psychophysiological studies (e.g., Greenwald et al., 1989; Lang et al., 1993). Its validity is further supported by its successful use as a dependable indicator of arousal in consumer studies (e.g., Alexander et al., 2015; Baldo et al., 2022; Langner et al., 2015; Maxian et al., 2013; Somervuo and Ravaja, 2013).

Second, the ability to continually measure EDA over time (Caruelle et al., 2019; Lajante & Ladhari, 2019) made it a perfect fit for our study objectives. Other approaches to capturing emotion over time involve either reporting how one feels at intervals (e.g., Dubé & Morgan, 1996, 1998; Maguire & Geiger, 2015) or inferring how others feel, for example, based on their speech/language (e.g., Verhoef et al., 2004), but reporting how one feels multiple times over time is both cumbersome and intrusive, and inferring how others feel over a long period may lack consistency and precision. We employed a psychophysiological measurement that avoids these challenges (Verhulst et al., 2019). Using sensors that record EDA values multiple times (4 Hz) in a second, we captured dynamic fluctuations in arousal over a service encounter, ensuring the appropriateness and relevance of our study design.

Measuring EDA involves measuring the skin’s electrical conductivity (Boucsein, 2012). When an emotion-eliciting stimulus is experienced, physiological changes happen in the body, including activation of the eccrine sweat glands, which briefly (for a few seconds) increases the skin’s electrical conductivity (Stern et al., 2000). This activation following an emotion-eliciting stimulus corresponds to the phasic component of EDA (known as phasic EDA), while tonic EDA corresponds to “background” activity (i.e., baseline level) and varies much more slowly, for instance, as a function of physical effort (Boucsein et al., 2012). The phasic EDA value, typically measured in microsiemens (µS), indicates the degree of arousal generated by the stimulus, with a greater phasic EDA value indicating higher arousal.

3.2. Procedure

Our procedure was identical in both study settings. First, the participants were equipped with a wireless EDA sensor (Empatica E4) and eye-tracking glasses ( Tobii Pro Glasses 2), enabling us to follow them remotely. They then completed an entrance survey. Next, a research assistant administered a stressor task to identify non-responders (those whose EDA does not respond to arousing stimuli).1 The participants then visited the store, receiving no instruction regarding the visit, browsing any aisle, and spending as much time as they wished in the store. In the hypermarket, a confederate dressed as a store employee offered a coupon to participants selected at random, an intervention that was unrelated to the present study. Finally, a research assistant met the participants at the store’s exit, and they filled out an exit survey including measures of customer approach response. Appendix A presents further procedural details, notably regarding the use of the EDA sensor.

3.3. EDA data analysis

We analyzed the EDA data following the steps recommended by Caruelle et al. (2019) and in accordance with the psychophysiology literature (e.g., Boucsein, 2012) and existing practices in the consumer literature (e.g., Baldo et al., 2022). We used continuous decomposition analysis in Ledalab V3.4.9 (Benedek & Kaernbach, 2010) to quantify phasic EDA values throughout each EDA recording.2 The continuous decomposition analysis consists in decomposing the EDA signal into its phasic and tonic components—tonic EDA corresponding to the baseline level of EDA and phasic EDA indicating the response to a stimulus (Boucsein et al., 2012).

3.4. Predictor variables

In the conceptual framework, we proposed to test the effect on customer approach response of the peak (H1), end (H2), trend (H3), and skewness of distribution of arousal levels (H4). We computed these variables as follows:

Peak. We computed the peak as the highest value of phasic EDA during the encounter and standardized it using Ben-Shakhar et al.’s (1975) formula to account for inter-individual differences in EDA values.

End. We computed the end as the average value of phasic EDA during the last third of the encounter and standardized it using the formula of Ben-Shakhar et al. (1975).3

Trend. We computed the trend by dividing the encounter into three periods of equal duration using Dubé and Morgan’s (1996) formula:

\[
\sum_{p=1}^{3} \frac{(A_{i(p+1)} - A_{i(p)}/A_{i(p)})}{(n - 1)}
\]

where \(A_{i(p)}\) is the average phasic EDA value for individual i in period p and n is the total number of periods (here, 3).

Skewness. We computed the skewness of the distribution of the phasic EDA values during the encounter using SKEW function in Microsoft Excel.

3.5. Outcome variables

Prior research (e.g., Andersson et al., 2012; De Nisco & Warnaby, 2014; Donovan & Rossiter, 1982; Sherman et al., 1997) suggests that customer approach response comprises various attitudinal and

1 Some pathologies (e.g., schizophrenia) impede electrodermal responsiveness to arousing stimuli (Boucsein, 2012), and a stressor task detects the lack of responsiveness. We adapted a mental arithmetic task commonly used to check electrodermal responsiveness (e.g., Blain et al., 2008; Poh et al., 2019). The participants were instructed to count backward by sevens from 700 and to say the numbers aloud to the research assistant.

2 Before that, we preprocessed the EDA data in Ledalab by applying a fifth-order low-pass Butterworth filter (1 Hz cutoff) to the EDA signal to reduce the high-frequency noise typically present in electrical signals (Boucsein, 2012). We then performed automated artifact detection using EDA Explorer (Taylor et al., 2015). We discarded all five-second windows identified by EDA Explorer as containing an artifact in accordance with Boucsein’s (2012) recommendations.

3 The peak-end rule does not indicate how long the end should be. Consequently, multiple computations of the end have emerged in studies testing this rule. In the study of Ariely and Loewenstein (2000), an episode is divided into three parts of equal duration: the start, middle, and final part. We found this partition relevant because it can be applied to episodes of any duration. We therefore decided to operationalize the end as the last third of the encounter.
behavioral responses. Thus, we used five operationalizations of customer approach response.

*Store visit duration.* Following Donovan and Rossiter (1982) and Sherman et al. (1997), we operationalized customer approach response as time spent in the store. Approach response is characterized by a desire to stay in and explore the environment, and greater time in the store indicates that response. Store visit duration was measured in second based on the time markers recorded by the EDA sensor.

*Customer spending.* Another operationalization of customer approach response is customer spending (Sherman et al., 1997). Because approach response is characterized by a desire to affiliate with the environment, greater spending indicates that response. The participants reported their spending in the local currency (Swedish crowns) in the exit survey.

*Unplanned purchasing.* Following Donovan and Rossiter (1982), who operationalized customer approach response as unplanned spending, we operationalized customer approach response as unplanned purchasing. Customer approach response is marked by a desire to explore the environment, which, in a store, is likely to result in more unplanned purchasing. We measured this with two items in the exit survey (Cronbach’s α = 0.630): “Today, I felt tempted to buy something that I had not planned to buy” and “During today’s visit, I totally held myself to the shopping list I had in mind” (reversed item). Each item was measured on a 7-point scale from “Strongly disagree” to “Strongly agree”.

*Satisfaction with the encounter.* Following Andersson et al. (2012) and Donovan and Rossiter (1982), who measured shopping enjoyment, and Wirtz et al. (2000) and Mattila and Wirtz (2001), who operationalized customer approach response as satisfaction with the service encounter, we measured satisfaction with the store visit. Approach response is characterized by a favorable attitude toward the environment, of which satisfaction is an indicator. Therefore, we measured satisfaction with the store visit in the exit survey, with the following items adapted from Eroglu and Macleit (1990): “To what extent are you satisfied or dissatisfied with today’s shopping trip at [store name]?” (1 = Dissatisfied, 7 = Satisfied); “How would you assess today’s shopping trip at [store name]?” (1 = Unpleasant, 7 = Pleasant; 1 = Unfavorable, 7 = Favorable; 1 = did not like it at all, 7 = I liked it very much). Eight participants had missing answers for one or several items on the satisfaction scale. These missing values were imputed using the person mean substitution approach (Downey & King, 1998; Huisman, 2000). The scale showed a good reliability (Cronbach’s α = 0.804).

*Future Intentions.* Following Donovan and Rossiter (1982), who operationalized customer approach response as intention to return to the store, and Jung and Namkung (2009), who operationalized it as behavioral intentions (to return, recommend, and say positive things), we measured future behavioral intentions toward the store in the exit survey. Because approach response is characterized by a desire to affiliate with and return to the environment, intentions to recommend and say positive things about the store and intentions to stay loyal to the store are indicators of an approach response in a store environment. To measure future intentions, we adapted the scale from Seiders et al. (2007) that consists of the following items: “How likely are you to recommend others to shop at [store name]?”; “How likely are you to say positive things about [store name] to other people?”; “How likely are you to shop more often at [store name]?”; and “How likely are you to continue shopping at [store name]?” (1 = Unlikely, 7 = Likely; Cronbach’s α = 0.735).

### 3.6. Participants

Of the 161 recruited participants, 36 were excluded. In the exit survey, the participants were asked to report any particularly positive or negative incident that they would have experienced in the store. Thirteen participants (hypermarket: 10; hardware store: 3) reported to have experienced a negative incident during the store visit. These participants were excluded from further analysis because the study of arousal on customer approach response requires that the situation or environment is not perceived as unpleasant. If a negative incident occurred, participants would exhibit an avoidance response. Twenty-three other participants were excluded because of technical problems with the equipment, problems with following the procedure, drop-out, or noise in the EDA data (for details, see Appendix B).

The resulting sample had 125 participants (hypermarket: 74; hardware store: 51). The average age was 42.79 years (SD = 16.73), and 37.6% were male; the rest were female. Of the participants, 45.6% had a shopping list, and 60.8% shopped with at least one other person. On average, they had been customers of the store for 8.87 years (SD = 7.03). The visits lasted 21.04 min on average (SD = 17.04) and resulted in a purchase by 88.8% of the participants. Out of the 74 participants in the hypermarket, 43 were offered a coupon.

### 4. Results

For each outcome variable, we performed a linear regression analysis with the peak, end, trend, and skewness of the distribution of arousal levels as predictor variables. In all five regression analyses, we controlled for mean arousal level, type of store (dummy variable coded 1 for hypermarket and 2 for hardware store), and offering a coupon (dummy variable coded 0 for no coupon and 1 otherwise).

In the regressions on customer spending and unplanned purchasing, we also controlled for having a shopping list (as reported in the entrance survey), because having one may reduce spending (Thomas & Garfand, 1993) and the likelihood of unplanned purchases (Inman et al., 2009). Similarly, in the regressions on satisfaction with the store visit and future intentions, we controlled for prior overall satisfaction with the store (measured on a 7-point scale prior to the store visit, in the entrance

---

**Table 2**

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>Store visit duration</th>
<th>Customer spending</th>
<th>Unplanned purchasing</th>
<th>Satisfaction</th>
<th>Future intentions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>7.462***</td>
<td>523.865***</td>
<td>1.992***</td>
<td>2.943***</td>
<td>3.871***</td>
</tr>
<tr>
<td>Mean arousal level</td>
<td>–1.500***</td>
<td>–275.900</td>
<td>0.539</td>
<td>–0.244</td>
<td>–0.156</td>
</tr>
<tr>
<td>Peak arousal level</td>
<td>0.143***</td>
<td>71.825***</td>
<td>0.266**</td>
<td>0.105*</td>
<td>0.038</td>
</tr>
<tr>
<td>End arousal level</td>
<td>0.089</td>
<td>14.231</td>
<td>0.075</td>
<td>0.171</td>
<td>0.012</td>
</tr>
<tr>
<td>Trend in arousal level</td>
<td>0.010</td>
<td>–0.096</td>
<td>–0.096</td>
<td>0.099</td>
<td>0.150</td>
</tr>
<tr>
<td>Skewness of the distribution of arousal values</td>
<td>–0.209***</td>
<td>–89.950*</td>
<td>–0.276</td>
<td>–0.238**</td>
<td>–0.094</td>
</tr>
<tr>
<td>Store</td>
<td>–0.559***</td>
<td>–326.377***</td>
<td>–0.076</td>
<td>0.491**</td>
<td>0.048</td>
</tr>
<tr>
<td>Coupon</td>
<td>0.061</td>
<td>173.687*</td>
<td>0.909**</td>
<td>0.360*</td>
<td>–0.074</td>
</tr>
<tr>
<td>Shopping list</td>
<td>223.201***</td>
<td>–0.017</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior overall satisfaction</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.560</td>
<td>0.342</td>
<td>0.035</td>
<td>0.156</td>
<td>0.207</td>
</tr>
</tbody>
</table>

*p < .1; ** p < .05; *** p < .01.

---

4 We computed the mean as the average phasic EDA value during the encounter, and we standardized it using the formula of Ben-Shahar et al. (1975) to account for inter-individual differences in EDA values.
survey), because these two outcomes are likely to be affected by prior cumulative satisfaction with the store (Bolton, 1998). Table 2 summarizes the results of the regression analyses.

First, we performed a log-transformation of the visit duration, as this variable was not normally distributed. The regression analysis on (log-transformed) visit duration uncovered a significant effect of the peak ($b = 0.143, p < .001$) and the skewness of the distribution of arousal levels ($b = -0.209, p = .001$). However, neither the end nor trend had a significant effect on (log-transformed) visit duration ($p > .1$ for both).

Second, regarding customer spending, we found a significant effect of peak arousal level ($b = 71.825, p = .012$) and a marginally significant effect of the skewness of the distribution of arousal levels ($b = -89.950, p = .089$). The end and trend had no significant effect on spending ($p > .1$ for both).

Third, unplanned purchasing was significantly predicted by peak arousal level ($b = 0.266, p = .040$), but no other variable (end, trend, or skewness) had a significant effect on unplanned purchasing ($p > .1$ for all three).

Fourth, we found that the skewness of the distribution of the arousal responses ($b = -0.238, p = .038$) had a significant effect on satisfaction with the encounter, while peak arousal level ($b = 0.105, p = .086$) had a marginally significant effect. The end and trend had no significant effect ($p > .1$ for both).

Fifth, none of the variables of interest had a significant effect on future intentions ($p > .1$ for the peak, end, trend, and skewness variables).

5. Discussion

First, it is worth noting that the mean arousal level had no significant effect on the variables operationalizing customer approach response (except store visit duration). This finding suggests that the mean level is little adequate in predicting approach response, which further suggests that variations in arousal level may be better predictors. The mean arousal level was included as a control variable.

Our overall results provide mixed support for our predictions (see Table 3 for a summary). Our prediction of the impact of peak arousal level on customer approach response (H1) received the most support, followed by our prediction regarding the impact of the positive skewness of the distribution of arousal levels (H4). Our other predictions (H2 and H3) were not empirically supported.

H1 was supported for all but one of the outcome variables. A greater peak arousal level significantly increased store visit duration, customer spending, and unplanned purchasing and marginally significantly increased satisfaction with the encounter. However, it did not have a significant effect on future intentions. Overall, this result suggests that the highest arousal level during an encounter has a positive effect on customer approach response, in line with our theorizing. Because the moment corresponding to the highest level of arousal indicates how emotionally engaging the experience can get, it is reasonable that it increases customer approach response.

H4 was supported for three of the five indicators of customer approach response: the positively skewed distribution of arousal levels significantly predicted both visit duration and satisfaction with the encounter and marginally significantly predicted customer spending. Importantly, this effect was negative on all three indicators. That is, although skewness significantly predicted a limited set of indicators, the consistent finding that this effect is negative is in line with the prediction in H4. This suggests that experiencing many moments of low arousal and relatively few of higher arousal (a positively skewed distribution) throughout a service encounter decreases customer approach response, all else being equal. We can interpret this result as follows: when the distribution of arousal levels is positively skewed, exciting moments (arousal levels at the upper end of the range) are rare but sufficient to make it salient that all the other moments of the encounter are little exciting (arousal levels at the lower end of the range), giving an overall impression of boredom, which naturally decreases customer approach response.

We found no empirical support for the predictions that arousal level at the end of the encounter (H2) and trend in arousal level (H3) impacted customer approach response. Fredrickson (2000) argues that a plausible explanation for the peak-end rule is that the peak and end are the most personally meaningful moments of an episode, notably because the end coincides with the achievement (or not) of a goal that one seeks to achieve in the given episode. Unlike other research contexts in which the end rule was demonstrated (e.g., childbirth; Chajut et al., 2014), the end of a store visit may not carry a strong sense of achievement, as purchase decisions are made throughout the visit, not only at the end. We speculate that for our participants, the end of the visit was not associated with personal meaning, which possibly explains the lack of significant effect on customer approach response of the end and its corollary, the trend.

H2 and H3 differ importantly from H1 and H4 in that the former hypotheses address the temporal aspect of variations in arousal level (i.e., when these variations occur during the encounter), whereas the latter do not. H1 and H4 propose, respectively, that the highest arousal level and the distribution of arousal levels predict approach response, independently of when in the encounter the high and low arousal levels are experienced. The findings thus suggest that variations in arousal level predict customer approach response but that the temporal position of these variations is not impactful.

Finally, of the five variables operationalizing customer approach response, one was not significantly predicted by any of the hypothesized features: future intentions. Even though approach responses encompass both present and future-oriented responses (Vieira, 2013), we found that variations in arousal level predicted the present approach response but not a future-oriented one, which may be explained by the participants having been customers of the stores for almost nine years on average. In such a long customer-firm relationship, every additional experience with the firm has little impact on future intentions compared to prior cumulative experience (Bolton, 1998). The fact that prior overall satisfaction significantly affected future intentions in our study suggests that this was the case.

Table 3
Summary of the findings.

<table>
<thead>
<tr>
<th>Positive effect of a greater peak arousal level (H1)</th>
<th>Positive effect of a greater end arousal level (H2)</th>
<th>Positive effect of an increasing trend in arousal level (H3)</th>
<th>Negative effect of a more positively skewed distribution of arousal levels (H4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer approach response</td>
<td>Customer spending</td>
<td>Unplanned purchasing</td>
<td>Satisfaction with the encounter</td>
</tr>
<tr>
<td>Store visit duration</td>
<td>Yes</td>
<td>Yes</td>
<td>Partial support</td>
</tr>
<tr>
<td>Customer spending</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Unplanned purchasing</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Future intentions</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
6. Conclusions

6.1. Theoretical and methodological contributions

Our research makes several theoretical and methodological contributions. First, it enriches the customer experience literature by studying how variations in arousal, which relates to the emotional dimension of customer experience, predict customer approach response in a service encounter. It is well established that customer experience is dynamic (Becker & Jaakkola, 2020; Kranzbühler et al., 2018; Lemon & Verhoef, 2016), but few studies capture the emotional dimension of customer experience over time. This is particularly true for studies examining arousal, which, with some exceptions (see Table 1), have captured it as a static phenomenon (measured at a single point in time). We add to this research stream by capturing arousal over time throughout the service encounter.

Thus, our research extends the existing knowledge on the predictive role of emotion (its arousal dimension) for customer outcomes in service encounters. It has long been recognized that emotion resulting from a service encounter matters (e.g., for satisfaction judgment), but our research demonstrates that how arousal unfolds during an encounter also matters. Customers’ arousal levels vary over time, and, importantly, how they vary (as captured by features such as the peak and skewness of the distribution) predicts customer approach response. Specifically, our study revealed that peak arousal level predicted four indicators of customer approach response and that the skewness of the distribution of arousal levels predicted three. In sum, our study of arousal over time provides insights that a summative assessment of arousal cannot deliver.

Furthermore, our research contributes to the psychology literature by jointly testing the model of evaluation by moments and range-frequency theory. To the best of our knowledge, and as can be observed in Table 1, ours is the first study to simultaneously test a prediction from range-frequency theory and predictions derived from the model of evaluation by moments. The prediction derived from range-frequency theory was supported (i.e., we found significant or marginally significant effects) for three of the five operationalizations of customer approach response. Regarding the predictions derived from the model of evaluation by moments, the prediction regarding peak arousal was largely supported, whereas the other two predictions (regarding the end and the trend) found no empirical support. Tully and Meyvis (2016) question the validity of the peak-end rule by demonstrating that the emotional state at the end of an episode does not systematically predict the evaluation of the episode. Our finding contributes to this research stream by documenting a context—store visits—in which arousal at the end of an episode did not significantly affect the evaluation of the episode.

The last contribution that our research makes is a methodological one by measuring arousal continuously and in situ throughout a service encounter using EDA measurement. De Keyser et al. (2015, 2021), Lemon and Verhoef (2016), and Morales et al. (2017) urge marketing researchers to study customer experiences using new types of data, including biometrics/physiological data, to obtain novel insights. Such data have been collected in laboratory settings (e.g., Boshoff, 2012; Verhulst et al., 2020), and our research adds to this research stream by collecting physiological (EDA) data in situ, measuring customers’ arousal levels continuously over time and preserving the richness of the stimuli experienced holistically at actual touchpoints. This method offers novel opportunities for customer experience research, but our study also reveals the challenges of the method, as technical problems with the equipment may necessitate removing participants from the final sample.

6.2. Managerial implications

Practitioners are well aware of the need to manage customer journeys to create great customer experiences (Boyarsky et al., 2016; Gibbons, 2019), but they often neglect the emotional aspect of those journeys, as Burns (2015) notes: “companies work hard to improve customer experience but often emphasize its utilitarian aspects of effectiveness and ease rather than emotion—how interactions make customers feel.” For some managers, customer experience management consists primarily of establishing touchpoints and connecting them seamlessly in the customer journey (Homburg et al., 2017). Our research invites practitioners to more closely consider customer emotion (in its arousal dimension) when managing customer experience in service encounters. Specifically, it suggests that practitioners should measure arousal throughout the encounter, not only at the end, to capture variations in customers’ arousal level over time and thus examine peak arousal level and the skewness of the arousal levels’ distribution to identify how to improve customer approach response.

Taken alone, our finding that peak arousal level increases customer approach response indicates that managers should design service encounters that make customers experience a highly arousing moment. With the mass-market device Fitbit now including an EDA sensor, it is plausible that, in the future, firms could collect EDA data directly from customers’ devices (with their consent) and intervene in real time, if necessary, to create such a highly arousing moment, using stimuli that have been preliminary identified as triggering high arousal.

However, our finding that a positively skewed distribution of arousal levels decreases customer approach response refines this recommendation. If most of the encounter triggers little to no arousal, one highly arousing moment, by being the only exciting one, will make the rest of the encounter seem even less emotionally engaging, thus reducing customer approach response. Instead, encounters should be designed to emotionally engage customers at moderate to high levels during most of the encounter. Infrequent moments of little or no arousal would then make it salient that most of the encounter is somewhat exciting, increasing customer approach response. The specific marketing actions needed to create arousing moments will depend on context, service type, and customers. To identify a relevant set of actions, we encourage firms to test how diverse stimuli (e.g., atmospherics; employees’ displayed emotions; unexpected in-store coupons) and their combinations impact customers’ arousal levels.

6.3. Limitations and directions for future research

Our research has the following limitations. First, arousal can have a positive or negative valence. To meaningfully interpret our results, we analyzed the data of participants who reported no negative incident during the encounter. An avenue for future research is to complement the EDA measurement method with a method that captures emotional valence (e.g., facial recognition). Researchers could thus verify that emotional arousal was positive without needing to survey the participants. Alternatively, a laboratory study using a predetermined script would give researchers greater control over the valence of the experience throughout the study.

Another avenue for future research related to valence is to apply our predictions to negative arousal. Some service encounters, such as hospital stays (McColl-Kennedy et al., 2017), and some touchpoints in the customer journey, such as service failure episodes (McColl-Kennedy et al., 2009), tend to generate negative emotion. In such instances, we would expect the features characterizing variations in arousal level (peak, end, trend, skewness) to have an effect in the opposite direction to that predicted in the present research. Because negative emotions are more impactful than positive ones (Baumeister et al., 2001), we would also expect the effect of these features to be stronger for negative than for positive arousal.

A second limitation relates to the test of predictions derived from the model of evaluation by moments. To compute the trend and the end, the store visit had to be divided into a certain number of periods, so we divided it into three equal periods for each participant: the start, middle, and end (Ariely & Loewenstein, 2000). However, customers may delimit their own store visits differently, with fewer or more periods or different
the missing markers accordingly. Taps on the button were too weak. We used the videos recorded by the eye-tracking glasses to identify when a phase started/ended and reconstituted by the participants, as it was done remotely through the eye-tracking software. Several time markers were missing in the EDA data, presumably because the research assistant paused the eye-tracking so as not to record the personal codes of participants paying with a bank card. The pause was unnoticeable.

The Empatica E4 can include time markers in the EDA data by simply tapping a button on the device. We used this feature throughout the study so as to break down the whole EDA recording into the study phases (see Fig. 1). Between each phase, a research assistant tapped the button to add a time marker signaling the end of one phase and the start of another. Furthermore, participants in our study had long relationships (almost nine years on average) with the stores, which may explain why we found no significant effect of variations in arousal level on future intentions. We encourage future research to replicate our study on customers with a short or no relationship with the service firm, as it would be insightful to determine whether variations in arousal level throughout a service encounter shape future intentions at an early stage of the customer-firm relationship.

Finally, it would be interesting to investigate how interruptions in the service encounter—interruptions that make customers perceive the encounter as partitioned—may attenuate the effect of variations in arousal level on customer approach response. Ariely and Zauberman (2003) have shown that individuals rely less on the trend in intensity to evaluate an episode when they perceive the episode to be partitioned rather than continuous. More research is needed to determine whether this finding applies to other features characterizing variations in arousal level. If so, we would expect that, in service encounters or customer journeys that customers perceive as partitioned (e.g., an IKEA visit comprising two parts: the visit to the store itself and a meal at the IKEA restaurant), the effect of variations in arousal level on customer approach response would be less.

CRediT authorship contribution statement

Delphine Caruelle: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. Poja Shams: Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Data curation, Conceptualization. Anders Gustafsson: Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Data curation, Conceptualization. Line Lervik-Olsen: Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The Knowledge Foundation (KK-stiftelsen), Sweden, funded this research.

Appendix A. Additional information regarding the procedure

We chose the Empatica E4 to record EDA because this sensor was developed for extended recording in ambulatory settings (Empatica, 2014). The Empatica E4 is designed as a wristband whose two electrodes for EDA measurement are placed on the inner wrist. The research assistants placed the wristband on the participants’ non-dominant hand, as recommended by Boucsein (2012).

The data measured by Empatica E4 were recorded by a dedicated mobile application from Empatica. When participants reached the cashier, a research assistant paused the eye-tracking so as not to record the personal codes of participants paying with a bank card. The pause was unnoticeable by the participants, as it was done remotely through the eye-tracking software.

The Empatica E4 can include time markers in the EDA data by simply tapping a button on the device. We used this feature throughout the study so that we could break down the whole EDA recording into the study phases (see Fig. 1). Between each phase, a research assistant tapped the button to add a time marker signaling the end of one phase and the start of another. Several time markers were missing in the EDA data, presumably because the taps on the button were too weak. We used the videos recorded by the eye-tracking glasses to identify when a phase started/ended and reconstituted the missing markers accordingly.

Appendix B. Exclusion of participants

<table>
<thead>
<tr>
<th>Reason for exclusion</th>
<th>Number of participants excluded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drop-out</td>
<td></td>
</tr>
<tr>
<td>The participant refused to perform the stressor task and dropped out of the study.</td>
<td>1 in the hypermarket</td>
</tr>
<tr>
<td>Procedure not strictly followed</td>
<td></td>
</tr>
<tr>
<td>• The participant went shopping at a different store after being equipped.</td>
<td>1 in the hypermarket</td>
</tr>
<tr>
<td>• The EDA recording device was placed on the participant’s dominant rather than non-dominant hand.</td>
<td>1 in the hypermarket</td>
</tr>
<tr>
<td>• The participant let a companion fill out the entrance and exit surveys without consulting them.</td>
<td>1 in the hypermarket</td>
</tr>
</tbody>
</table>

(continued on next page)
The EDA recording was incomplete, as the EDA recording device turned off unexpectedly (e.g., due to a spent battery).

In some EDA recordings, one or several time markers were missing, presumably because the taps on the device’s button were too weak. Because these markers are necessary to identify the exact start and end of the visit, we used videos recorded by the eye-tracking glasses to locate these missing markers. However, for 10 participants in the hypermarket, we were unable to locate the marker signaling the end of the visit, as the eye-tracking recording was paused then. Without knowledge of when the visit ended, it was impossible to compute metrics related to variations in arousal throughout the visit.

Noise in EDA data. Artifact detection by EDA Explorer led to the removal of all EDA data during the first, second, or last (third) part of the store visit, making it impossible to compute the trend metric.

No electrodermal responsiveness. Three participants had no phasic EDA value above 0.05 μS during the visit. When EDA is recorded in field studies, an appropriate threshold for the phasic EDA value to qualify as a skin conductance response is 0.05 μS (Greene-Klein & Baun, 2001; Greene-Klein, 2005). The lack of response during the visit indicates either that the visit was not emotionally arousing or that the participants were non-responders. We checked the phasic EDA values of the three participants during the stressor task and found that, for all of them, no phasic EDA value during the stressor task reached the threshold of 0.05 μS. This indicates a strong likelihood that these participants were non-responders.

+ Among which one participant was also excluded for reporting a negative incident.

References


Benedek, M., & Baun, 2001; Groeppel-Klein, 2005). The lack of response during the visit indicates either that the visit was not emotionally arousing or that the participants were non-responders. We checked the phasic EDA values of the three participants during the stressor task and found that, for all of them, no phasic EDA value during the stressor task reached the threshold of 0.05 μS. This indicates a strong likelihood that these participants were non-responders.


