An emotional rollercoaster: Electrophysiological evidence of emotional engagement during a rollercoaster ride with and without a Virtual Reality add-on

Bastiaansen, Marcel; Oosterholt, Monique; Mitas, Ondrej; Han, Dai-In; Lub, Xander

published in
Journal of Hospitality and Tourism Research
2022

DOI (link to publisher)
10.1177/1096348020944436

document version
Publisher's PDF, also known as Version of record

AN EMOTIONAL ROLLER COASTER: ELECTROPHYSIOLOGICAL EVIDENCE OF EMOTIONAL ENGAGEMENT DURING A ROLLER-COASTER RIDE WITH VIRTUAL REALITY ADD-ON

Marcel Bastiaansen
Breda University of Applied Sciences
Tilburg University
Monique Oosterholt
Ondrej Mitas
Danny Han
Breda University of Applied Sciences
Xander Lub
Breda University of Applied Sciences
Nyenrode Business University

Emotions are crucial ingredients of meaningful and memorable tourism experiences. Research methods borrowed from experimental psychology are prime candidates for quantifying emotions while experiences are unfolding. The present article empirically evaluates the methodological feasibility and usefulness of ambulatory recordings of skin conductance responses (SCRs) during a tourism experience. We recorded SCRs in participants while they experienced a roller-coaster ride with or without a virtual reality (VR) headset. Ride elements were identified that related to physical aspects (such as accelerations and braking), to events in the VR environment, and to the physical theming of the roller coaster. VR rides were evaluated more positively than normal rides. SCR time series were meaningfully related to the different ride elements. SCR signals did not significantly predict overall evaluations of the ride. We conclude that psychophysiological measurements are a new avenue for understanding how hospitality, tourism and leisure experiences dynamically develop over time.

KEYWORDS: *emotions; roller coaster; experience; virtual reality; skin conductance*

INTRODUCTION

There is increasing consensus among tourism, hospitality, and leisure scholars that emotions are crucial ingredients for meaningful and memorable experiences (Bastiaansen et al., 2019; Moyle et al., 2017; Skavronskaya et al., 2017;

Journal of Hospitality & Tourism Research, Vol. 46, No. 1, January 2022, 29–54
DOI: 10.1177/1096348020944436

Zajchowski et al., 2016). At the same time, it has repeatedly been emphasized that studying emotions using traditional research methods that rely on self-report is limited by substantial validity issues (e.g., Larsen & Fredrickson, 1999; Mauss & Robinson, 2009). These developments have prompted scholars in our field to search for alternative, less biased ways of measuring the emotions that are felt during leisure and tourism experiences. Research methods borrowed from experimental psychology and psychophysiology have been proposed as prime candidates for quantifying emotions while experiences are unfolding. These methods offer the additional benefit of fine-grained temporal resolution (Bastiaansen et al., 2019; Li et al., 2015; Li et al., 2018; Moyle et al., 2017; Tröndle et al., 2014). With these recent developments in mind, the present article aims to address empirically the potential usefulness of ambulatory recordings of skin conductance as an objective measure of emotional engagement during real-life leisure, tourism, and hospitality experiences.

Theme parks and specifically roller coasters provide an ideal platform for studying emotional engagement in a controlled and staged experience identical for all visitors. For one, a roller-coaster ride is a tightly staged experience. It has identical experiential elements for all visitors. These elements follow a strict and uniform timing. Second, virtual reality (VR) roller coasters, in which visitors are given the choice to ride the roller coaster with or without a VR headset, constitute a well-controlled quasi-experimental manipulation of the actual experience. With a VR add-on, visitors are immersed into a 3D animated world, which adds ride elements to the roller-coaster experience that do not exist in the physical world. Combining this with the fact that riding a roller coaster in an amusement park is a very natural tourism activity to engage in, a VR coaster as an experience platform combines full ecological validity with an almost perfectly controlled quasi-experimental design.

This first and primary contribution this study aims to make is a methodological one: to establish the feasibility of using state-of-the-art psychophysiological research methods (in particular, skin conductance measurements) to measure emotional engagement as it unfolds over time in ecologically valid, real-life tourism and leisure settings.

Second, by exploiting the fact that skin conductance is measured continuously over time, we address hypotheses based on peak–end theory (Fredrickson, 2000) about the relation between the timing of emotional engagement and overall evaluations of the experience. This knowledge is not only relevant for academic research, but may also open doors to managers in this industry to measure and optimally design experiences for their guests.

LITERATURE REVIEW

Lived and Remembered Experiences

Experiences play a central role in the tourism, hospitality, and leisure economy (Pine & Gilmore, 1998, 2011). Consumers pursue experiences that they consider to be meaningful and memorable. For researchers that want to study those experiences, this implies that experience measures have to be available that are both reliable and valid. Therefore, tourism, hospitality, and leisure scholars have attempted to define the experience construct and to understand how experiences are being evaluated and remembered (e.g., Bastiaansen et al., 2019; Pearce & Zare, 2017; Scott & Le, 2017).

Tourism, hospitality, and leisure scholars are increasingly aware of the fact that emotions play an important role in shaping experiences (Bastiaansen et al., 2019; Moyle et al., 2017; Skavronskaya et al., 2017; Zajchowski et al., 2016). This awareness is grounded in more fundamental psychological work on experiences (Jantzen, 2013) and on work of Kahneman (2011; Kahneman et al., 1993; Kahneman & Tversky, 2003). Notably, peak–end theory (Fredrickson, 2000), which posits a clear relation between how lived emotions transform into remembered experience, has attracted much attention, both academically and professionally. Whether an experience is memorable is important to guiding future behavioral decisions, such as repeat visits and willingness to recommend (Kahneman, 2011; Zajchowski et al., 2016). In brief, peak–end theory states that the emotional peak in an experience, and the emotions felt toward the end of it, are essential in how the experience is remembered. However, peak–end theory has been formulated on the basis of experimental paradigms which induce quite homogeneous and unidimensional experiences, often in a medical context, such as experiencing pain during a colonoscopy or while immersing a hand in ice-cold water. It is therefore an open question whether peak–end theory also applies to the more heterogenous and complex experiences found in hospitality, tourism, and leisure. Initial evidence based on self-report measures of emotional valence and arousal (Strijbosch et al., 2019) suggests that during the leisure experience of watching a short VR movie, the average emotional valence and arousal, rather than peaks and ends, are the best predictors of overall evaluations. These findings challenge the applicability of peak–end theory for hospitality, tourism and leisure experiences. In this study, we further address peak–end theory by verifying to what extent peoples' overall evaluations of a roller coaster ride can be predicted by the peak emotional engagement during a roller coaster ride, and the emotional engagement at the end of the ride.

Challenges for Measuring Experiences

The theoretical developments and empirical questions sketched above pose methodological challenges for studying tourism, hospitality, and leisure experiences. Traditional approaches to studying experiences (e.g., Brown & Novak-Leonard, 2013; J.-H. Kim et al., 2012; Lee & Smith, 2015) ignore the ebb and flow of emotional engagement during experiences. Measurement tools that are sufficiently sensitive to register time-varying changes in emotions with sufficient temporal resolution are needed. This requires an understanding of emotions and how they are expressed.

There is general consensus among emotion researchers that emotions are a response to stimuli that are seen as personally relevant. This response is expressed at three levels: the phenomenological, physiological, and behavioral levels (Ekman, 2016; Mauss & Robinson, 2009). It follows that emotions can be measured at each of three levels of expression: phenomenological responses can be assessed through self-report, physiological responses through electrophysiological recordings, and behavioral responses through observation (Bastiaansen et al., 2019; Mauss & Robinson, 2009).

Electrophysiological measures are a promising tool for studying experiences, for several reasons. First, such measures have an excellent time resolution (typically less than a second) and therefore can track the temporal dynamics of emotional engagement as an experiential episode unfolds. Second, a substantial body of psychophysiological literature is available that documents how electrophysiological measures relate emotional valence and arousal—at least in carefully controlled laboratory settings. Finally, over the past few years wearable recording devices have become available that allow for recording electrophysiological measures continuously while people are freely engaging in leisure activities which makes it possible to achieve high levels of ecological validity.

Using Skin Conductance to Measure Experiences in Real Time

In the present study, we aim to establish whether rapid changes in skin conductance are useful for studying tourism, hospitality, and leisure experiences. Skin conductance signals consist of slow changes in skin conductance level (SCL), which are mainly related to thermoregulatory processes. Superimposed on SCL are brief, phasic changes (phasic skin conductance responses [SCRs]) which are related to emotional arousal. By applying the proper mathematical procedures to skin conductance measurements (continuous deconvolution; Benedek & Kaernbach, 2010a, 2010b), SCRs align well with the moment that an emotion is experienced, although they typically lag behind the emotion-triggering event by 2 to 3 seconds (Benedek & Kaernbach, 2010b). In carefully controlled psychophysiological laboratory experiments, SCRs to discrete (i.e., temporally isolated) stimuli are substantially larger following emotionally salient stimuli as compared to emotionally neutral stimuli (for a very comprehensive review of skin conductance methodology and skin conductance research, see Boucsein, 2012). Note however that SCRs do not discriminate between positive and negative emotional responses (Boucsein, 2012).

These developments suggest that SCRs can be used to quantify the ebb and flow of emotional engagement during tourism, hospitality, and leisure experiences. At the same time, one should take care to realize that moving from controlled laboratory experiments using highly sensitive equipment and simple, temporally isolated stimuli, to continuous SC measurements with wearable devices in complex real-life situations is a substantial step that, however attractive, should be taken cautiously. Loss of data quality and experimental control may be so impactful that the endeavor is challenging at best. Two issues are particularly relevant. First, when skin conductance sensors move relative to the skin (which is likely to happen in real-life situations), this results in so-called motion artifacts, which are high peaks in the recorded signal that do not reflect true skin conductance. Therefore, in real-life settings care should be taken to detect, and where possible, remove those artifacts from the recorded signal (Chen et al., 2015; Taylor et al., 2015).

Second, real-life situations offer little experimental control. As visitors behave freely, it is difficult to determine exactly what stimuli they are experiencing at any given moment in time. It is therefore even more difficult to relate subsecond changes in SCR measurements to that experience. Therefore, tightly staged experiences, which offer only small variations in the lived experience across individuals, may be a good starting point for establishing the usefulness of physiological recordings in our field.

Despite these caveats, researchers in our field are beginning to use skin conductance as a tool for measuring experiences in real-life settings (for reviews, see Bastiaansen et al., 2019; Li et al., 2014; Scott et al., 2017). A number of studies have addressed the emotional engagement of tourists as they engage in a city trip (J. Kim & Fesenmaier, 2015) or city walk (Birenboim et al., 2019; Shoval et al., 2018). Although the initial results are promising, these studies did not attempt to relate continuous measures of emotional engagement to overall evaluations of the experience, or other outcome measures. Another set of studies has addressed visitor's emotional engagement during a museum visit (Tröndle et al., 2014; Tröndle & Tschacher, 2012; Tschacher et al., 2012, and carefully described how different patterns of emotional engagement are observed in different types of visitors. However, none of the published studies to date have addressed important methodological issues such as motion artifacts or decomposing the skin conductance signal into tonic (SCL) and phasic (SCR) components. Also, in all the studies described above, the measured experience was very loosely staged, leading to substantial heterogeneity in what different participants were actually experiencing. This heterogeneity makes it challenging to relate measures of skin conductance to both the lived and the remembered experience. In the present study, we aim to overcome these shortcomings in three different ways: by engaging participants in a tightly staged experience (a roller-coaster ride), by using state-of-the-art analysis techniques (motion artifact correction, continuous deconvolution), and by relating SCRs to a specific outcome measure (overall evaluation of the experience).

The Present Study

In the present study, we compare the experience of a roller coaster ride with and without a VR add-on. The use of VR has been proposed as a highly promising tool in the design and optimization of experiences (Han, 2019). In the context of theme parks, VR has recently been used to enhance or redefine roller-coaster rides (Jung et al., 2018). Overlaying a virtual environment on a physical roller-coaster ride allows for a reinterpretation of the physical sensations and for creating different scenarios in the virtual setting. Due to the resulting economic benefits, VR coasters in theme parks are an area of fast development in VR implementation (Baker, 2016).

This form of entertainment provides a level of immersion that is so far exclusive to VR (Williams & Mascioni, 2017). According to Siegrist et al. (2019, the level of experienced immersion depends on how closely the virtually generated environment reflects the real physical environment. This immersion is quantified by as the sense of *presence*—the subjective experience of being in one place while being physically located in another (Witmer & Singer, 1998). A sense of presence in the virtual environment can detach the user from the physical world. Kim and Biocca (1997) describe this phenomenon with measures of *arrival* and *departure*, where arrival refers to the sense of being in one's physical location, while departure refers to the sense of being somewhere else. In sum, presence is a relevant measure in determining to what extent a user is immersed in a VR environment.

In our view, a VR coaster is an ideal context for studying the usefulness of SCR for experience measurement. Notably, a VR coaster combines three strong points: First, as the ride is fixed, both in terms of duration and in terms of timing and temporal order of the different ride elements. Thus, a VR coaster provides a tightly staged experience, and hence very clear temporal information about which stimuli participants experience at which moment. Second, a roller-coaster ride in a theme park setting has perfect ecological validity. And third, the VR add-on allows for a very natural yet well-controlled quasi-experimental manipulation in the ride experience, creating two highly similar conditions for which SCRs can be contrasted. One potential disadvantage of the VR coaster as an experience platform, however, is that the physical movements of the coaster may induce motion artifacts in the SC data.

The study took place in Europapark, a theme park located in the southwest of Germany. Europapark features a roller-coaster attraction named Alpenexpress Enzian. The ride has a 100-second fixed duration, and is a moderate-intensity ride, without loopings, spirals, or highly intense turns and speed changes. Without VR add-on, the ride is lightly themed, with traditional Austrian landscapes (Alpine mountain peaks and meadows), while with the VR add-on visitors experience views as if they were virtually flying an airplane. Supplement Figure 1 (available online) gives a visual impression of the roller-coaster ride and the VR add-on.

Based on a careful inspection of the actual ride and of the VR environment, we identified three different types of ride elements in the roller-coaster experience (see Methods for details). Physical ride elements, commonly experienced by visitors in the VR and in the non-VR (henceforth NVR) ride; themed ride elements specific to the NVR ride; and third, themed ride elements specific to the VR ride.

We measured skin conductance of participants during the (VR or NVR) ride with wristbands. After the ride, participants filled out a questionnaire, asking about the sense of presence they experienced and about their overall evaluation of the ride. With these data, we aim to address the following research questions:

> **Research Question 1:** Is the roller-coaster ride with VR as immersive as the one without VR?

Presence, as a proxy of the immersiveness of an experience, has mainly been studied in the context of VR experiences. To verify the extent to which visitors taking the VR ride were fully immersed into the virtual environment, we performed a direct comparison between the levels of presence experienced during the VR and NVR rides.

> **Research Question 2:** Does the VR add-on lead to a better evaluation of the roller-coaster ride?

VR coasters are increasingly popular in theme parks and are seen as a promising tool for enhancing and/or redefine roller-coaster experiences. A direct comparison between how VR and NVR rides are evaluated in terms of emotional valence and arousal clarified whether a VR add-on improved the experience.

> **Research Question 3:** Can a continuous measure of emotional engagement during the ride be reliably related to the different elements of the roller-coaster ride?

The main purpose of the present study is to determine whether state-of-the-art physiological research methods measure emotional engagement as it unfolds over time in ecologically valid, real-life tourism and leisure settings. We address the three following subquestions:

> **Research Question 3.1:** Are the temporal profiles of SCRs for the VR and NVR rides different from one another?

If the VR theming of the roller-coaster ride influences the levels of emotional engagement that visitors experience during the ride, this difference should be substantial.

> **Research Question 3.2:** Are changes in SCRs over time related to the ride elements that make up the roller-coaster experience?
> **Research Question 3.3:** Which type of ride element (physical vs. VR-themed or NVR-themed) is most strongly correlated with SCR measures?
> **Research Question 4:** Can overall evaluations of the ride be predicted from emotional engagement during the ride?
> **Research Question 4.1:** From peak–end theory, it follows that peaks and ends in SCRs are good predictors of overall ride evaluations.

In contrast, a recent proposal has been that for more heterogeneous tourism experiences, average emotional engagement (operationalized in the present study as average SCRs across the entire ride) is a better predictor of overall evaluations. We address this question by comparing how well peak, end, peak-and-end SCRs, and average SCRs predict overall ride evaluations.

> **Research Question 4.2:** Alternatively, it is conceivable that the SCRs during specific ride elements that make up the experience predict overall evaluations better than theory-informed peak–end predictors.

We address this question by establishing how well SCRs during the different ride elements predicts overall evaluations of the ride, and by comparing the outcomes to those of research question 4.1.

METHOD

Participants

Eighty-one participants took part in the experiment. Participants were 18 years or older (M = 32.4, SD = 9.34). Of the 81 participants, 34 chose to ride the roller coaster without VR add-on (NVR group), and 47 chose either the Ed Euromaus (N = 16) or Sky Explorers (N = 31) VR ride. Note that participants were not randomly assigned to one of the rides, and therefore our research design is quasi-experimental rather than experimental in the strict sense. As the group of participants that opted for the Ed Euromaus VR ride was very small, these participants were excluded from further analysis. Of the remaining 65 participants, the physiological data of 12 participants turned out to be either missing due to failure of the equipment, or devalued by excessive motion artifacts (see Data Analysis section). These participants were excluded from all further analyses.

The final sample thus consisted of 53 participants. Of these, 29 were in the NVR group, and 24 in the VR group. The demographic characteristics for both groups, and their previous experience with VR, are provided in the online Supplement Table 1.

Stimulus Materials

The roller-coaster ride experience was designed differently for the NVR and VR groups. A short description of the two is provided below. The timing of the different ride elements is given in the online Supplement Table 3.

NVR Ride

During the NVR ride, the participants encountered several lightly themed areas. The theming of the area around the roller coaster was in an Austrian Alpine style, which was reflected in the overall storyline and in the theming of the ride itself. The ride started outside after leaving the station. There the participants could spot a few Alpine forest animals such as bears. They continued their journey inside a cave, themed as a mine. The participants encountered a specific smell and could notice in the first part that the cave had dimmed lighting and bright diamonds. The second part of the cave was dark, and the speed of the roller coaster increased in this final turn. The cave was followed by short section outside, before the coaster rushed through the station and began a second round. The ride ended at a slow pace outside before it entered the station for the second time, where it finished.

VR Ride

Participants who opted for the VR ride could choose one of two possible virtual environments: Ed Euromaus or Sky Explorers. As Ed Euromaus was excluded from the study, only the Sky Explorers ride is described below. During the ride the movement of the VR experience was synchronized with the roller coaster layout, to prevent nausea and other negative effects. The participants were exposed to a mediated 360°-computer-generated environment, meaning they were able to move their head in all directions without disturbing the story or the environment.

In the Sky Explorers VR storyline, the participants were virtually sitting in an airplane. The experience was based on another Europapark attraction called "Voletarium." The participant's plane followed another airplane and immediately left the room to go outside. The plane passed underneath a bridge with a train track, and through several landscapes. The plane entered a small village and flew through several streets. After flying over a square, the plane did a sideways looping to fit through a door that opened just in time. The plane then left the village and flew through a waterfall into a cave. The timing of entering the cave corresponded to entering the mine (NVR ride). For a moment, the plane flew above a big lake in the cave and above a huge fish. Then it left the cave and flew above a rocky, snowy landscape with goats before making its final turn and landing.

Data Collection

Procedure. Data were collected over a 3-day period in May 2018 in Europapark, Germany. Participants were approached at the entrance of the attraction. Every third person was approached, regardless of whether they chose the NVR or the VR ride. Participants were explained the research goal and what they could expect of participation. After verbal agreement, they were further informed through written instructions in their own language (English, German, French, or Dutch). After having given written informed consent, participants were asked to fill out a 2-minute questionnaire, containing demographic background questions. Subsequently, before they joined the queue for the attraction, an Empatica E4 wristband (a wearable electrophysiological recording device) was put on the wrist of the participants.

In order to minimize potential motion artifacts in the electrophysiological recordings, participants were instructed to relax the arm on which the wristband was attached, to put the corresponding hand in their lap, and not to clench the bar of the roller-coaster cart during the ride. Then the participants joined the queue and rode the roller coaster.

After the ride, participants were approached at the exit. They returned the Empatica wristband, and they filled out the second part of the questionnaire. After filling out this 3- to 5-minute questionnaire, one question (see the section postride questionnaire for details) was asked verbally, and the answer was recorded on a mobile device.

Both the preride and the postride questionnaire were available in four different languages, based on validated translations: English, German, French, and Dutch. Participants were asked what their preferred language was and were handed the proper questionnaire accordingly.

Preride Questionnaire. All participants filled out the same preride questionnaire, regardless of the type of ride they chose. The questionnaire included questions about age, gender, which roller-coaster experience they were going to ride, if they had previous experience with this roller coaster and if they had any previous experiences with VR. This last question was only applicable for participants from the VR group.

Physiological data. These data were recorded with the Empatica E4 wearable wristband. These commercially available wristbands have been shown to record physiological signals (skin conductance, heart rate, skin temperature) with decent levels of scientific quality (Birenboim et al., 2019; Ollander et al., 2016; Ragot et al., 2017). In addition, the wristband also records acceleration data. Skin conductance and accelerometer data were continuously sampled at a built-in frequency of 4 Hz and stored on the Empatica device for further offline processing. Measurement of physiological responses started at the moment the participants received the Empatica E4 and lasted until they were taken off by the experimenter after the ride. The duration of the roller coaster was exactly 1:40 minutes (100 seconds). The collected physiological recordings varied in length between 5 to 30 minutes, depending on the length of the queue. To obtain an indicative time alignment between the physiological recordings and onset/offset of a ride, the experimenter held a separate Empatica wristband on which time stamps were recorded at the onset and offset of the ride. Accelerometer data were collected from the Empatica wristbands to synchronize physiological recordings with the on- and offsets of the ride with subsecond precision (see section on Data Analysis).

Postride Questionnaire. The postride questionnaire contained three sections. It included questions on presence, on discrete emotions experienced during the ride, and on an overall evaluation of the ride.

In the first section, measurement of presence was based on the ITC–Sense of Presence Inventory questionnaire, which is validated across a range of media experiences, including VR (Lessiter et al., 2001). A subset of the items of the ITC–Sense of Presence Inventory questionnaire was selected by taking the items that loaded highest on each determinant of mediated presence in the study of Lessiter et al. (2001). This resulted in a measurement scale with the following subscales (see Supplement Table 3, available online): sense of physical space (three items), engagement (three items), ecological validity (three items) and negative effects (four items). Responses were collected on a five-point Likert scale ranging from *strongly disagree* to *strongly agree*.

The second section addressed to which extent eleven discrete emotions were experienced during the ride. These data were not analyzed in the present study.

Third, participants were asked to evaluate the ride on three different aspects: an overall evaluation, an evaluation of the peak experience during the ride, and an evaluation of the end of the ride. Evaluations for these three aspects were given on two dimensions, valence and arousal. Valence was prompted by the questions "How positive or negative did you feel during the ride/during the most exciting moment of the ride/toward the end of ride?" However, responses were collected on a 5-point scale ranging from *very negative* to *very positive*. Arousal was prompted by the question "How calm or excited did you feel during the ride/during the most exciting moment of the ride/toward the end of ride?" and responses were collected on a 5-point scale ranging from *calm* to *excited*.

Data Analysis

Postexperience Questionnaire. For the presence scale, averages of the subscales were computed, and a reliability analysis was performed for each subscale. In addition, differences in presence between the VR and NVR rides were evaluated with independent-samples *t* tests for each subscale.

The overall evaluation measures of valence and arousal were computed, and independent-samples *t* test were used to compare overall valence and overall arousal between the VR and the NVR groups. The results of this analysis address our second research question.

Skin conductance data processing. The skin conductance and accelerometer data were extracted from the Empatica wristbands and imported into MATLAB for further analysis. Processing and analysis of these data was performed using a set of MATLAB functions developed by the authors that is available in open source on request. First, skin conductance data were precisely time-synchronized with the onset and offset of the ride. For this, we took advantage of the fact that the ride starts with an initial acceleration, which was readily identifiable from the Empatica accelerometer data. Skin conductance segments of 120 seconds were then extracted from the recordings, corresponding to a 10-second preride interval, the 100-second ride, and a 10-second postride interval.

Skin conductance data, especially when collected through wearable devices, can be contaminated by motion artifacts (e.g., Taylor et al., 2015; for review, see Boucsein, 2012). Motion artifacts result from pressure on the device, or from movement of the sensors contained in the device relative to the skin. They take the shape of high-amplitude, short-lived spikes (typically less than 1 second, see also Figure 1), and as such can be distinguished from true SCRs, which have lower amplitude and, crucially, last for several seconds (Boucsein, 2012). Although more elaborate methods for detecting and removing motion artifacts from skin conductance data have been devised (e.g., Chen et al., 2015; Kelsey et al., 2017; Taylor et al., 2015), none of them are publicly available. We therefore developed a simple, supervised method for detecting and correcting the skin conductance signal for motion artifacts. Artifacts are detected by applying a z transform to a moving time window (here 10 seconds) and visualizing the signal in that time window whenever a z value exceeds a set threshold (here $\pm$ 3). The experimenter then decides whether or not the detected peak or trough is a motion artifact and should be corrected. In case a motion artifact is clearly identified, it is removed from the signal by linearly interpolating the signal from the left-hand border of the spike to its right-hand border. In the present data, about half of the 53 participants' data (after excluding 12 participants that contained excessive artifacts, see also the Participants section) did not contain any clear motion artifacts, while in the other half, between 1 and 10 artifacts were typically removed following this procedure. In case of ambiguity (e.g., the signal takes the shape of a motion artifact but lasts longer than 1 second), the signal was not altered to avoid removing true SCRs from the data. Note that this careful approach may nevertheless leave some motion artifacts behind in the data.

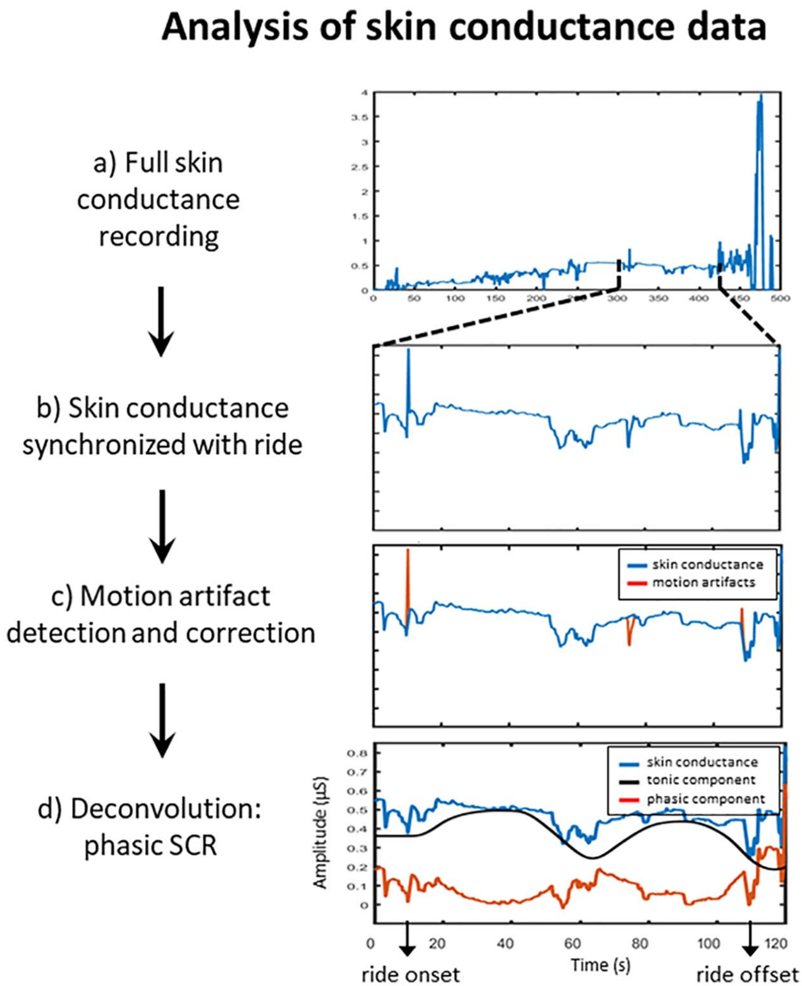
In a next step, each participant's skin conductance data were subjected to a continuous deconvolution (Benedek & Kaernbach, 2010a), which splits the signal into a tonic component and a phasic component. As stated in the literature review, the phasic component consists of a superposition of SCRs, and is most closely related to emotional arousal. We will refer to these signals as SCRs in the remainder of this article. The open-source MATLAB toolbox Ledalab (Benedek & Kaernbach, 2010a) was used for this analysis step.

Finally, SCR signals for each participant were used as a basis for the statistical analysis. In addition, SCRs were averaged across participants separately for the NVR and the VR rides for display purposes. The entire signal analysis procedure is visualized in Figure 1.

SCR differences between the groups. To address Research Question 3.1, we first tested whether the SCR time courses differed between the two groups. An independent-samples t test was conducted for each data point during the 5-second preride interval and the 100-second ride, amounting to 420 t tests. False discovery rate correction (Benjamini & Hochberg, 1995) was used to correct for multiple comparisons.

Modeling ride elements. For Research Questions 3.2 and 3.3, we sought to characterize the two rides by defining the most important ride elements. The

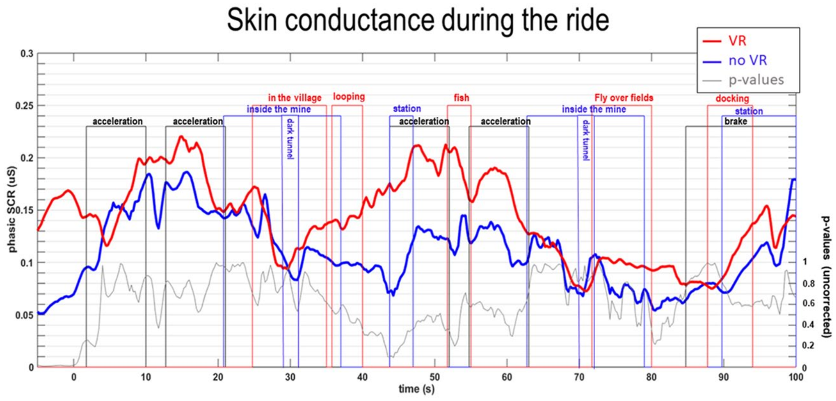
Figure 1
Illustration of the Data Analysis Procedure for Skin Conductance, Using Example Data From One Participant



Note: (a) The entire skin conductance recording includes queueing, the ride itself, and a postride interval. (b) A 120-second segment is taken from the recording, precisely synchronized with the ride. (c) Motion artifacts are detected and corrected. (d) Continuous deconvolution separates the skin conductance in a tonic component and a phasic component. The latter is used in the subsequent statistical analyses.

following criteria were established for defining ride elements: they should be sufficiently long-lasting (minimally 3 seconds), be clearly identifiable in time (clear onset and offset) and represent conceptually distinct "events" or "periods" of the ride (although admittedly, the latter criterion is somewhat arbitrary). We

Figure 2
Grand Average Phasic Skin Conductance for the Virtual Reality (VR) Group (N = 24) and the Nonvirtual Reality Group (N = 29) During the Roller-Coaster Ride



Note: t = 0 corresponds to ride onset. Boxcar functions indicate the timing of the physical ride elements common to both groups (black boxcar), and of the thematic ride elements separately for the VR group (red boxcar) and no VR group (blue boxcar). Gray line in the figure indicates uncorrected p values (right-hand y axis) from the independent-samples t tests at each sampling point.

distinguished between physical ride elements, which were commonly experienced by both groups, and themed ride elements, which were different for the NVR and VR groups. Physical ride elements comprised four successive accelerations, and a period of braking at the end of the ride. Themed ride elements for the NVR group were "entering the mine," "dark tunnel," and "station," all of which were present twice during the ride (once for each round of the roller coaster). For the VR group, themed ride elements were "flying above a fish," "entering the village," "virtual looping," "flying over a plain with fields and rocks," and "docking the plane." The different ride elements, with their respective timings, are given in the online Supplement Table 2, and graphically represented as boxcar functions in Figure 2.

The three boxcar functions from Figure 2 were taken as simplified models of the different rides. To establish whether SCRs vary as a function of the ride, these boxcars (taking values of either zero or one) were correlated with the SCR data. The following general analysis strategy was adopted: one or more boxcar functions (for details see below) were used as a predictor(s) in a linear regression analysis, with individual SCR level as the dependent variable. Note that correlations among boxcars were moderate (physical—VR: $r = -.33$; physical—NVR: $r = -.27$; VR–NVR: $r = .37$). Therefore, multicollinearity was not considered to be an issue in the multiple regressions. Regression analyses were then performed for each participant separately, in which time points served as observations. A single-sample t test then verified whether standardized beta weights (for each of the predictor variables) across participants were significantly different from zero. The average R^2 across participants was computed as an estimate of the effect size.

Using this strategy, the following analyses were performed:

- Predicting SCRs of all participants from the boxcar function modeling physical ride elements
- Predicting SCRs of the VR participants from the boxcar function modeling the VR ride elements
- Predicting SCRs of the VR participants from the boxcar function modeling the physical ride elements and the one modeling the VR ride elements
- Predicting SCRs of the NVR participants from the boxcar function modeling the NVR ride elements
- Predicting SCRs of the NVR participants from the boxcar function modeling the physical ride elements and the one modeling the NVR ride elements

Predicting overall evaluations from peak, end, peak–end, and average SCR. To address research question 4.1, SCR values were considered in a 105-second interval, starting 5 seconds before the onset of the ride to the end of the ride. First the peak SCR was identified, and the average SCR in a 5-second window around that peak was extracted (PEAK measure). In addition, the average SCR during the past 5 seconds of the ride was computed for each participant (END measure). Furthermore, these two measures were combined (averaged) to create a PEAK-AND-END measure. Finally, the average SCR was computed for the entire 105-second interval (AVERAGE measure). The PEAK, END, PEAK_AND_END and AVERAGE variables were then used as independent variables in simple linear regression analyses to predict overall self-reported evaluation measures (overall valence and overall arousal) from the postride questionnaire (see Strijbosch et al., 2019, for a similar approach).

Predicting overall evaluations from SCRs during the ride elements. For each ride element, the average of the SCR values in the time window of that ride element was computed for each participant separately. Variables thus obtained were then used in multiple linear regression analyses to predict overall self-reported evaluation measures (overall valence and overall arousal) from the postride questionnaire. The following regression analyses were conducted:

- Predicting overall valence from the average SCRs corresponding to the five physical ride elements (all participants).
- Predicting overall arousal from the average SCRs corresponding to the five physical ride elements (all participants).
- Predicting overall valence from the average SCRs corresponding to the five physical and the 6 NVR ride elements (NVR participants).
- Predicting overall arousal from the average SCRs corresponding to the five physical and the 6 NVR ride elements (NVR participants).

- Predicting overall valence from the average SCRs corresponding to the five physical and the 5 VR ride elements (VR participants).
- Predicting overall arousal from the average SCRs corresponding to the five physical and the 5 VR ride elements (NVR participants).

RESULTS

Presence

A reliability analysis on the four subscales of the presence scale indicated that the subscales of Sense of Physical Space (Cronbach's α = .84), Ecological Validity (Cronbach's α = .81), and Negative Effects (Cronbach's α = .85) were highly reliable. The responses to the individual items of those subscales were therefore averaged for each subscale and used for subsequent analysis. For the subscale Engagement, reliability across the three items was low (Cronbach's α = .38). After removal of the item "I felt involved (in the environment)," Cronbach's alpha went up to .73. Therefore, for this subscale the average was computed for the remaining two items only. The scores on the resulting presence subscales are presented in the online Supplement Figure 2.

As online Supplement Figure 2 suggests, scores on the subscale Sense of Physical Space were significantly larger for the VR group ($t_{49} = 2.75, p = .008$). No significant differences between the two groups were observed for the other subscales (Engagement: $t_{49} = .87, p = .391$; Ecological Validity: $t_{48} = .18, p = .858$; Negative Effects: $t_{51} = .29, p = .758$).

Overall Evaluations

Overall valence for the VR ride ($M = 4.32$, standard error (SE) = 0.153) was larger than for the NVR ride ($M = 3.79, SE = 0.152$), $t_{49} = 2.39, p = .021$. Also, overall arousal for the VR ride ($M = 3.77, SE = 0.146$) was larger than for the NVR ride ($M = 2.48, SE = 0.202$), $t_{49} = 4.86, p < .001$. The results are graphically represented in the online Supplement Figure 3.

Skin Conductance Data

Grand averages of SCRs, separated by the VR and NVR groups, are presented in Figure 2. The Figure suggests that SCRs are overall larger for the VR ride than for the NVR ride, and the temporal evolution of the signal appears to roughly follow the different ride elements. Initially, during the 5-second preride interval SCRs are substantially larger for the VR than for the NVR group. Furthermore, for both groups SCRs increase during the successive accelerations, and during the braking toward the end of the ride. In addition, a marked dip in SCRs is observed for both groups at the time of the "dark tunnel." In the middle section of the ride, SCRs in the VR group appears to be markedly larger than for the NVR group. However, these visual impressions are further qualified by the outcomes of the statistical analyses.

Table 1
Regressions Predicting SCR From Boxcar-Modelled Ride Elements

Participants	Predictors	Mean R^2 Across Participants	t Test on β Values Across Participants
All	Physical ride elements only	.045	t_{52} = 3.19, p = .002**
VR group	VR ride elements only	.048	t_{23} = −3.08, p = .005**
VR group	Physical ride elements and VR ride elements	.085	t_{23} =, 1,76, p = .091; t_{23} = −2.53, p = .019*
NVR group	NVR ride elements only	.055	t_{28} = −0.86, p = .399
NVR group	Physical ride elements and NVR ride elements	.097	t_{28} = 1.99, p = .056; t_{28} = −0.26, p = .797

Note: SCR = skin conductance responses; VR = virtual reality; NVR = nonvirtual reality.
*$p < .05$. **$p < .01$.

SCR differences between the groups. Averaged across the entire ride, SCRs are not different between the two groups ($t_{52} = 0.62$, $p = .538$). Furthermore, the timepoint-by-timepoint Independent-samples t tests of the differences in SCR levels between the two groups (see the gray line in Figure 2) indicate that SCRs only differ statistically between the groups in the 5-second preride interval, where SCR levels are significantly higher for the VR group than for the NVR group.

Correlating ride elements with SCR. The results of the regression analyses that relate ride elements to SCR signals are summarized in Table 1. The averages of the R^2 values across participants indicate that the boxcar functions modelling the different ride elements account for between 4.5% and 9.7% of the variance in SCRs. Across all participants, physical ride elements yield per-participant beta values that are consistently higher than zero, indicating that the physical ride elements consistently (i.e., across participants) account for variance in SCRs. The same holds for the VR ride elements in the VR group. For the NVR group however, the NVR ride elements do not consistently account for variance in SCRs.

Predicting overall evaluations from SCRs: Peak, end, peak–end, and average SCRs. The regression analyses shown in Table 2 were aimed at predicting the overall, postexperience evaluations of the ride (both in terms of overall valence and of overall arousal) from the peak, end, combined peak–end, and average SCR levels. R^2 values are generally low, indicating that less than 4% of the variance in overall evaluations can be predicted from these regressors. In addition, the regression models were far from reaching significance, indicating that overall evaluations cannot be predicted from these regressors.

Predicting overall evaluations from SCR during the different ride elements. The regression analyses shown in Table 2 were aimed at predicting the postexperience evaluations of the ride from the average of the SCR values in the time

Table 2
Regressions Predicting Overall Evaluations From Peak, End, Peak–End, and Average SCR Across the Entire Ride

Predictor	Dependent variable	R^2	*F* statistic	*p*
Peak	Valence	.003	$F_{1, 49} = 0.14$	.701
	Arousal	.037	$F_{1, 49} = 1.92$	.171
End	Valence	.005	$F_{1, 49} = 0.24$	.629
	Arousal	.019	$F_{1, 49} = 0.93$	.341
Peak–end	Valence	.004	$F_{1, 49} = 0.81$	.666
	Arousal	.032	$F_{1, 49} = 1.66$	.203
Average	Valence	.007	$F_{1, 49} = 0.35$	.556
	Arousal	.039	$F_{1, 49} = 1.96$	.167

Note: SCR = skin conductance responses.

windows of the different ride elements. R^2 values are substantial, indicating that as much as 43% of the variance in overall evaluations is accounted for when SCR during physical and VR ride elements were used to predict overall evaluations. However, all the regression models were far from reaching significance, indicating that overall evaluations cannot be predicted from these regressors.

DISCUSSION

The present article addresses two issues. First, it constitutes an attempt to establish the feasibility of using SCRs for studying the temporal profile of emotional engagement during tourism, hospitality, and leisure experiences. Second, the article addresses hypotheses based on peak–end theory (Fredrickson, 2000) about the relation between (the timing of) emotional engagement on the one hand, and overall evaluations of the experience on the other. SCRs were measured while participants rode a roller coaster with (VR) or without (NVR) a VR add-on. In addition, self-report measures of presence and overall evaluations of the roller-coaster ride were collected.

Levels of presence were largely similar for the VR and NVR rides, with the exception that the VR ride yielded higher scores on the dimension "sense of physical space." Furthermore, the VR ride was evaluated more positively and as being more arousing compared with the NVR ride. SCR data showed a temporal profile that could be both meaningfully and significantly related to the different elements of the roller-coaster ride. SCR data did not significantly predict post-experience evaluations of the ride, however.

Self-Reports Indicate That a VR Add-on Makes for a Better Experience

Presence was experienced in similar ways for the VR and NVR rides, for the dimensions Engagement, Ecological Validity, and Negative Effects. Surprisingly, the dimension Sense of Physical Space showed even larger scores for the VR ride compared to the NVR ride. These data answer our first research question (*Is the roller-coaster ride with VR as immersive as the one without VR?*), and further confirm that spectacular levels of immersion can be obtained through VR (Williams & Mascioni, 2017). With levels of presence that are comparable between real-life and VR environments, our data even suggest that VR is as immersive as real life.

Postexperience evaluation measures further indicated that participants felt more positive and more aroused during the VR ride than during the NVR ride. Together, the self-report data clearly indicate that the VR add-on led to a better experience of the roller-coaster ride. Thus, our data confirm suggestions that have been made in the literature regarding the experience-enhancing nature of VR (Huang et al., 2016; Wei et al., 2019).

Skin Conductance Responses Do Not Differentiate Between VR and NVR Rides

There are observable differences in SCRs between the VR and NVR rides (Figure 2). SCRs tend to be of higher magnitude during the VR ride, especially during the accelerations and braking. This seems to be an empirical confirmation of the suggestion that a VR add-on "can trick users into thinking drops, launches, twists and turns are in fact more extreme than they are in actuality" (Burt & Louw, 2019, p. 185). Also, there is a substantial difference in SCR between the rides in the middle portion, between roughly 35 seconds and 60 seconds after ride onset. This coincides with the VR ride elements Looping and Flying over a big fish. This seems to suggest that these VR ride elements are particularly emotionally engaging.

However, the two SCR profiles are not significantly different from each other, except for the 5 seconds interval before the onset of the ride, where the SCR is significantly higher for the VR ride than for the NVR ride. The initial difference in the preride segment may be attributed to the fact that, for the VR ride, the experience has already started (they are already in the virtual environment), while the NVR participants are merely waiting for the ride to begin.

Skin Conductance Responses Align With Different Ride Elements

Visual inspection of the temporal profile of SCRs suggests a clear correspondence of the increases and decreases in SCRs with the different elements of the roller-coaster ride. In both types of ride, SCRs increased during the four consecutive accelerations and during the braking at the end. In addition, when ride elements were modelled as boxcar functions, they significantly predicted between 4.5% and 9.7% of the variance in the SCR signal, at least for physical and VR ride elements (Table 1). The boxcar-modelled NVR ride elements did not correlate with the SCR signal. In sum, although effect sizes are small overall, we observed the strongest relationship between SCRs and physical ride elements, a weaker relationship with VR ride elements, and no significant relationship with NVR elements. Therefore, our data clearly indicate that changes in SCRs can be meaningfully related to the different elements of the roller-coaster ride.

In addressing Research Question 3.3 (*Which type of ride element [physical vs. VR-themed or NVR-themed--> is most strongly correlated with SCR measures?*), it is of interest that the physical ride elements relate most strongly with SCR measures, and that this holds both for the group of participants that did the VR ride as for the NVR participants. It seems then that the physical ride elements are a stronger driver of the SCR signal than the nonphysical, themed elements.

Skin Conductance Responses Do Not Predict Overall Evaluations

Our results very clearly show that SCR data do not predict evaluations. Both when predictors are based on peak–end theory and when predictors are based on a model of the different ride elements, regression models are far from being significant and have very low predictive power. Thus, in our data, overall evaluations of the ride cannot be predicted from SCR measurements. Therefore, our results do not differentiate between peak–end theory and possible competing alternatives (as proposed e.g., by Strijbosch et al., 2019).

To our knowledge, the present study is the first to address whether in-the-moment, physiological measurements can predict self-reported experience outcomes such as overall evaluations of emotional valence and emotional arousal. The total absence of predictive value is surprising, given the putative role of emotional engagement in shaping experiences (Bastiaansen et al., 2019; Moyle et al., 2017; Skavronskaya et al., 2017; Zajchowski et al., 2016). In our view there are two possible explanations for this null finding. One explanation would be related to technical issues such as a poor signal-to-noise ratio and a resulting lack of statistical power (see the section on limitations below). On a more conceptual or theoretical note, it could be that physiological data such as skin conductance reflect qualitatively different emotion processes compared to the self-response data that constitute overall evaluations. Mauss and Robinson (2009) have described in careful detail that different techniques for measuring emotions share only little common variance, and that each technique may tap into different aspects of how emotions arise and are perceived (see also Barrett et al., 2007 for a more theoretical perspective on this). One proposal, which is close to Kahneman's notion of two parallel decision making systems (Kahneman, 2011; Kahneman & Riis, 2005), is that physiological measurements reflect more automatic, unconscious emotional processes, whereas self-report items ask for a conscious and more rational evaluation of what has been felt and experienced. It is difficult in our view to devise sound empirical validations of such a proposal. In the context of the current study we thus adhere to Mauss and Robinson's

(2009) conclusion that "experiential, physiological and behavioral measures are all relevant to understanding emotion and cannot be assumed to be interchangeable" (p. 209).

LIMITATIONS

In addition to the observed lack of predictive power of SCR measures in predicting overall evaluations of the experience, the lack of significant differences between the SCR signals during the VR and NVR rides is remarkable given the robust correspondence of SCRs with ride elements. One explanation for the observed null findings may be the relatively high noise levels in the SCR data. An argument in favor of such an explanation is the high R^2 values observed in some of the analyses (e.g., Supplement Table 4, available online), which indicate that up to 43% of the variance in evaluations are predicted by SCR data. This oddly contrasts with the very high p-values that are observed in the same analyses.

As discussed in the literature review, moving from recordings with very sensitive equipment in well-controlled lab settings to using wearable devices in field settings is a substantial step, leading to attenuated signal-to-noise ratio (SNR) in the measurements, and by consequence to reduced statistical power. Although we took care to remove motion artifacts whenever they could be identified clearly, it is likely that the data contain residual artifacts. A roller-coaster ride comprises an adversarial context for data collection in this respect. Motion artifacts are likely to occur in any ambulatory measurement of skin conductance. In our view, therefore, subsequent studies using skin conductance measurements in real-life tourism, hospitality, and leisure experiences should be aware of the low SNR in their data, and the resulting loss of statistical power. We see two ways to deal with this effectively. One is by using relatively large numbers of study participants. The current sample (53 participants, divided over two groups), is larger than the typical sample of $N = 20$ to 25 used in laboratory studies of psychophysiological measures, but it appears to be on the lower border of what is needed in order to achieve reasonable SNR in ambulatory measurements. Another way of increasing SNR is by further improving signal processing techniques, notably artifact detection and correction techniques. We have used a simple, self-developed motion artifact correction technique (see Method), but more advanced techniques have been proposed, and should become more widely available (Chen et al., 2015; Kelsey et al., 2017; Taylor et al., 2015). Our study shows that using state-of-the-art data analysis techniques is indispensable for reliably and validly using skin conductance as a tool for measuring emotional engagement during experiences.

A second limitation is that relating subsecond changes in SCRs to moment-by-moment lived experience is only meaningful if SCRs are averaged across several participants. This assumes that the experience itself is highly comparable across these participants. With experiences that are heterogeneous across participants (e.g., a free walk in a theme park), the relationship between (averaged) SCRs and the experience itself becomes less obvious. We therefore feel that the current approach is best suited for studying tightly staged experiences that allow for little individual variation in the experience.

Finally, a limitation of the current study is that participants were not randomly assigned to the VR or NVR rides, but that participants self-selected which ride they experienced. This leaves the possibility that observed differences between rides are due to nonobserved differences in the two samples, such as differences in novelty-seeking or other psychological parameters, rather than due to differences in the ride itself.

PRACTICAL IMPLICATIONS AND CONTRIBUTION TO KNOWLEDGE

The present study has addressed the feasibility of using ambulatory physiological recordings to study the time course of emotional engagement during a tourism experience as it unfolds. Despite the acknowledged limitations of the study, our results suggest that skin conductance recordings, when carefully analyzed; can be meaningfully related to experiencing a roller-coaster ride. Future studies should further address methodological issues such as improving SNR in ambulatory recordings, and address issues of statistical power. If such methodological issues are overcome, the present approach may be a good candidate for further studying the relationship between emotions and experience, and address hypotheses derived from peak–end theory and competing alternative models. In addition, future studies should further explore the intricate relationship between physiological measures and postexperience self-report measures.

On a more practical note, given the observed correspondence between SCRs and elements that make up the roller-coaster ride, the approach we used—if further improved—holds promise as a tool to assist industry professionals in optimizing their customers' experiences. The detailed time information that SCR measures yield is potentially valuable in identifying which touchpoints in a visitor journey emotionally engage visitors, and which touchpoints do not. This information would allow for evidence-based optimization and redesign of visitor journeys. Similarly, SCR time information could potentially aid the design of VR environments for roller coasters, and of (tightly staged) tourism experiences more generally.

CONCLUSION

In conclusion, our study demonstrates that skin conductance, given proper analysis and further methodological development, has potential as a tool for measuring emotional engagement during tightly staged tourism, hospitality, and leisure experiences. Skin conductance signals are unfettered by response biases and have high temporal resolution. These advantages are relevant for academics studying the relationship between emotions and experience, as well as for managers and designers in the industry. With continuous technical improvements, measuring SCRs may potentially develop into a new and useful tool for developing, optimizing, evaluating and validating the design of tourism, hospitality, and leisure experiences.

ORCID iDs

Marcel Bastiaansen https://orcid.org/0000-0003-2865-7859
Danny Han https://orcid.org/0000-0001-7839-6760

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

Baker, C. A. (2016). Creative choices and fan practices in the transformation of theme park space. *Transformative Works*, *22*. https://doi.org/10.3983/twc.2016.0974

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology*, *58*, 373-403. https://doi.org/10.1146/annurev.psych.58.110405.085709

Bastiaansen, M., Lub, X., Mitas, O., Jung, T., Passos Acenção, M., Han, H., Moilanen, T., Smit, B., & Strijbosch, W. (2019). Emotions as core building blocks of an experience. *International Journal of Contemporary Hospitality Management*, *31*(2), 651-668. https://doi.org/10.1108/IJCHM-11-2017-0761

Benedek, M., & Kaernbach, C. (2010a). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80-91. https://doi.org/10.1016/j.jneumeth.2010.04.028

Benedek, M., & Kaernbach, C. (2010b). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, *47*(4), 647-658. https://doi.org/10.1111/j.1469-8986.2009.00972.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B*, *57*(1), 289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Birenboim, A., Dijst, M., Scheepers, F. E., Poelman, M. P., & Helbich, M. (2019). Wearables and location tracking technologies for mental-state sensing in outdoor environments. *The Professional Geographer*, *71*(3), 449-461. https://doi.org/10.1080/00330124.2018.1547978

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.). Springer. https://doi.org/10.1007/978-1-4614-1126-0

Brown, A. S., & Novak-Leonard, J. L. (2013). Measuring the intrinsic impacts of arts attendance. *Cultural Trends*, *22*(3-4), 223-233. https://doi.org/10.1080/09548963.2013.817654

Burt, M., & Louw, C. (2019). Virtual reality enhanced roller coasters and the future of entertainment–audience expectations. *World Leisure Journal*, *61*(3), 183-199. https://doi.org/10.1080/16078055.2019.1639274

Chen, W., Jaques, N., Taylor, S., Sano, A., Fedor, S., & Picard, R. W. (2015, August 25-29). *Wavelet-based motion artifact removal for electrodermal activity* [Paper presentation]. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy.

Ekman, P. (2016). What scientists who study emotion agree about. *Perspectives on Psychological Science*, *11*(1), 31-34. https://doi.org/10.1177/1745691615596992

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, *14*(4), 577-606. https://doi.org/10.1080/026999300402808

Han, D.-I. D. (2019). Calling for user-centric VR design research in hospitality and tourism. *Hospitality & Society*, *9*(2), 237-246. https://doi.org/10.1386/hosp.9.2.237_7

Huang, Y. C., Backman, K. F., Backman, S. J., & Chang, L. L. (2016). Exploring the implications of virtual reality technology in tourism marketing: An integrated research framework. *International Journal of Tourism Research*, *18*(2), 116-128. https://doi.org/10.1002/jtr.2038

Jantzen, C. (2013). Experiencing and experiences: A psychological framework. In J. Sundbo & F. Sørensen (Eds.), *Handbook on the experience economy* (pp. 146-170). Edward Elgar. https://doi.org/10.4337/9781781004227.00013

Jung, T., tom Dieck, M. C., Rauschnabel, P., Ascenção, M., Tuominen, P., & Moilanen, T. (2018). Functional, hedonic or social? Exploring antecedents and consequences of virtual reality rollercoaster usage. In T. Jung & M. C. tom Dieck (Eds.), *Augmented reality and virtual reality* (pp. 247-258). Springer. https://doi.org/10.1007/978-3-319-64027-3_17

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, *4*(6), 401-405. https://doi.org/10.1111/j.1467-9280.1993.tb00589.x

Kahneman, D., & Riis, J. (2005). Living, and thinking about it: Two perspectives on life. In F. A. Huppert, N. Baylis, & B. Keverne (Eds.), *The science of well-being* (pp. 285-304). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198567523.003.0011

Kahneman, D., & Tversky, A. (2003). Experienced utility and objective happiness: A moment-based approach. In I. Brocas & J. D. Carrillo (Eds.), *The psychology of economic decisions: Volume 1: Rationality and well-being* (pp. 187-208). Oxford University Press

Kelsey, M., Palumbo, R. V., Urbaneja, A., Akcakaya, M., Huang, J., Kleckner, I. R., Barrett, L. F., Quigley, K., Sejdic, E., & Goodwin, M. S. (2017, May 5). *Artifact detection in electrodermal activity using sparse recovery* [Paper presentation]. Compressive Sensing VI: From Diverse Modalities to Big Data Analytics. https://doi.org/10.1117/12.2264027

Kim, J., & Fesenmaier, D. R. (2015). Measuring emotions in real time: Implications for tourism experience design. *Journal of Travel Research*, *54*(4), 419-429. https://doi.org/10.1177/0047287514550100

Kim, J.-H., Ritchie, J. B., & McCormick, B. (2012). Development of a scale to measure memorable tourism experiences. *Journal of Travel Research*, *51*(1), 12-25. https://doi.org/10.1177/0047287510385467

Kim, T., & Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-mediated Communication*, *3*(2), JCMC325.

Larsen, R. J., & Fredrickson, B. L. (1999). Measurement issues in emotion research. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 40-60). Russel Sage Foundation.

Lee, H. M., & Smith, S. L. (2015). A visitor experience scale: historic sites and museums. *Journal of China Tourism Research*, *11*(3), 255-277. https://doi.org/10.1080/19388160.2015.1083499

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments*, *10*(3), 282-297. https://doi.org/10.1162/105474601300343612

Li, S., Scott, N., & Walters, G. (2015). Current and potential methods for measuring emotion in tourism experiences: A review. *Current issues in Tourism*, *18*(9), 805-827. https://doi.org/10.1080/13683500.2014.975679

Li, S., Walters, G., Packer, J., & Scott, N. (2018). A comparative analysis of self-report and psychophysiological measures of emotion in the context of tourism advertising. *Journal of Travel Research*, *57*(8), 1078-1092. https://doi.org/10.1177/0047287517733555

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion 23*(2), 209-237. https://doi.org/10.1080/02699930802204677

Moyle, B. D., Moyle, C.-l., Bec, A., & Scott, N. (2017). The next frontier in tourism emotion research. *Current Issues in Tourism*, *22*(12), 1393-1399. https://doi.org/10.1080/13683500.2017.1388770

Ollander, S., Godin, C., Campagne, A., & Charbonnier, S. (2016, October 9-12). *A comparison of wearable and stationary sensors for stress detection* [Paper presentation]. 2016 IEEE International Conference on Systems, Man, and Cybernetics. https://doi.org/10.1109/SMC.2016.7844917

Pearce, P. L., & Zare, S. (2017). The orchestra model as the basis for teaching tourism experience design. *Journal of Hospitality and Tourism Management*, *30*(March), 55-64. https://doi.org/10.1016/j.jhtm.2017.01.004

Pine, B. J., & Gilmore, J. H. (1998). Welcome to the experience economy. *Harvard Business Review*, *76*(4), 97-105.

Pine, B. J., & Gilmore, J. H. (2011). *The experience economy*: Harvard Business Press.

Ragot, M., Martin, N., Em, S., Pallamin, N., & Diverrez, J.-M. (2017, June 14). *Emotion recognition using physiological signals: Laboratory vs. wearable sensors* [Paper presentation]. International Conference on Applied Human Factors and Ergonomics. https://doi.org/10.1007/978-3-319-60639-2_2

Scott, N., Gao, J., & Ma, J. Y. (2017). *Visitor experience design* (Vol. 5). CABI. https://doi.org/10.1079/9781786391896.0000

Scott, N., & Le, D. (2017). Tourism experience: A review. In N. Scott, J. Gao, & J. Y. Ma (Eds.), *Visitor experience design* (Vol. 5). CABI. https://doi.org/10.1079/9781786391896.0000

Shoval, N., Schvimer, Y., & Tamir, M. (2018). Real-time measurement of tourists' objective and subjective emotions in time and space. *Journal of Travel Research*, *57*(1), 3-16. https://doi.org/10.1177/0047287517691155

Siegrist, M., Ung, C. Y., Zank, M., Marinello, M., Kunz, A., Hartmann, C., & Menozzi, M. (2019). Consumers' food selection behaviors in three-dimensional (3D) virtual reality. *Food Research International*, *117*, 50-59.

Skavronskaya, L., Scott, N., Moyle, B., Le, D., Hadinejad, A., Zhang, R., Gardiner, S., Coghlan, A., & Shakeela, A. (2017). Cognitive psychology and tourism research: state of the art. *Tourism Review*, *72*(2), 221-237. https://doi.org/10.1108/TR-03-2017-0041

Strijbosch, W., Mitas, O., van Gisbergen, M., Doicaru, M., Gelissen, J., & Bastiaansen, M. (2019). From experience to memory: On the robustness of the peak-and-end-rule for complex, heterogeneous experiences. *Frontiers in Psychology*, *10*, Article 1705. https://doi.org/10.3389/fpsyg.2019.01705

Taylor, S., Jaques, N., Chen, W., Fedor, S., Sano, A., & Picard, R. (2015, August 25-29). *Automatic identification of artifacts in electrodermal activity data* [Paper presentation]. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. https://doi.org/10.1109/EMBC.2015.7318762

Tröndle, M., Greenwood, S., Kirchberg, V., & Tschacher, W. (2014). An integrative and comprehensive methodology for studying aesthetic experience in the field: Merging movement tracking, physiology, and psychological data. *Environment and Behavior*, *46*(1), 102-135. https://doi.org/10.1177/0013916512453839

Tröndle, M., & Tschacher, W. (2012). The physiology of phenomenology: The effects of artworks. *Empirical Studies of the Arts*, *30*(1), 75-113. https://doi.org/10.2190/EM.30.1.g

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K., & Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96-103. https://doi.org/10.1037/a0023845

Wei, W., Qi, R., & Zhang, L. (2019). Effects of virtual reality on theme park visitors' experience and behaviors: A presence perspective. *Tourism Management*, *71*(April), 282-293. https://doi.org/10.1016/j.tourman.2018.10.024

Williams, K., & Mascioni, M. (2017). *The out-of-home immersive entertainment frontier: Expanding interactive boundaries in leisure facilities*. Routledge. https://doi.org/10.1016/j.tourman.2018.10.024

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225-240.

Zajchowski, C. A., Schwab, K. A., & Dustin, D. L. (2016). The experiencing self and the remembering self: Implications for leisure science. *Leisure Sciences*, *39*(6), 561-556. https://doi.org/10.1080/01490400.2016.1209140

Submitted January 30, 2019
Accepted June 11, 2020
Refereed Anonymously

Marcel Bastiaansen (e-mail: bastiaansen4.m@buas.nl) is a senior lecturer for Academy for Leisure & Events at the Breda University of Applied Sciences, Breda, The Netherlands. **Monique Oosterholt** (e-mail: oosterholtmonique@gmail.com) is a former MSc student at the Breda University of Applied Sciences, Breda, The Netherlands. **Ondrej Mitas** (e-mail: mitas.o@nhtv.nl) is a senior lecturer at the Breda University of Applied Sciences, Breda, The Netherlands. **Danny Han** (e-mail: han.d@nhtv.nl) is a senior lecturer for Academy of Hotel and Facility Management, at the Breda University of Applied Sciences, Breda, The Netherlands. **Xander Lub** (e-mail: x.lub@nyenrode.nl) is a former associate professor at the Breda University of Applied Sciences, Breda, The Netherlands.