Tilburg University

More is not better

Mitas, Ondrej; Mitasova, H.; Millar, G; Boode, Wilco; Neveu, V; Hover, Moniek; van den Eijnden, F.; Bastiaansen, Marcel

Link to publication in Tilburg University Research Portal

More is not better: The emotional dynamics of an excellent experience

Emotions are a crucial mechanism of overall evaluation in tourism experiences. Current tourist experience models do not explain how temporal dynamics of emotion add up to overall experience evaluations, however. We analyzed 15 participants' experiences at the Vincentre visitor center and guided outdoor tour in the village of Nuenen, the Netherlands. This Vincent van Gogh-themed experience led to a wide range of intent to recommend and emotional engagement, measured as continuous phasic skin conductance, across participants and exhibits. We conducted mixed-effects models of emotional engagement as a function of proximity to exhibits and intent to recommend. The most positively evaluated experiences featured ebbs and flows of emotional engagement, not one continuous 'high,' with more emotion during the middle phases of the experience. These findings suggest that tourist experience models should account for a complex relationship between emotions experienced and overall evaluation. Simply put, more emotion is not always better.

Introduction

Emotions are crucial to tourism experiences. It has been broadly recognized that emotions embody the value in tourism experiences (Knobloch, Robertson, & Aitken, 2017), are important mediators of recommendation and other target outcomes (Hosany & Gilbert, 2010; Prayag, Hosany, & Odeh, 2013), and fuel the social meanings of tourism experiences (Mitas, Yarnal, & Chick, 2012). Emotions also give practitioners the best entry point for measuring and managing experiences (Bastiaansen et al., 2019; Li, Scott, & Walters, 2015; Moyle, Moyle, Bec, & Scott, 2017; Skavronskaya et al., 2017).

The relationships among emotions and overall evaluations are controversial, however. Most cross-sectional studies show that positive emotions during a tourism experience predict positive overall evaluations and intent to recommend (Hosany & Gilbert, 2010; Prayag et al., 2013), although substantial variation is left unexplained. Furthermore, some studies have shown that a context-appropriate combination of positive *and* negative emotions was associated with positive overall evaluation (Nawijn & Fricke, 2015). In other words, it is not clear if simply *more* emotion during a tourism experience is *better*. Furthermore, cross-sectional studies leave the role of temporal dynamics unexplained (Bastiaansen et al., 2019). In the present paper, we report on an experience measurement of a complex tourism experience, comprised of a museum visit and village tour, to address the relationship between temporal and spatial dynamics of experience and overall experience evaluation. We situate this study in the context of existing tourism and psychology research on experiences and their outcomes.

Literature review

There is an extensive conceptual and empirical literature on the topic of tourist experiences. In a thorough and insightful review, Scott and Le (2017), conclude that theoretical models in this literature are generally constrained by the discipline from which they originate. Thus, there no coherent understanding of tourism experience is possible based on this literature. Furthermore, Scott and Le state that the psychological mechanisms of desired experience outcomes are a 'black box,' so largely unknown. In the present literature review, we cite recent works which posit that emotions are a key mechanism of experiential outcomes. Then, we review two models of emotion development over time, showing that neither fully answers how emotion development over a tourist experience adds up to an overall evaluation.

In synthesizing and extending extant tourist experience research, Scott and Le discuss concepts from cognitive psychology such as attention, emotion, memory, appraisal, involvement, immersion, and cognitive absorption. Pearce and Zare (2017) similarly presented the orchestra model of experience, wherein sensory, cognitive, affective, and behavioral dimensions interact to form an experience. Both models feature emotion prominently, though it is only one of several psychological mechanisms discussed.

Subsequently, Bastiaansen et al. (2019) proposed the Breda model of tourist experiences, which posits that emotion is the most important of these psychological mechanisms. In brief, their model posits that the mind cuts the continuous, stream-of-consciousness sense of experiencing something into discrete episodes based on mental models of temporal units, such as 'lunch,' 'visit,' 'weekend,' or 'vacation.' For example, in asking participants to recall their day one episode at a time, the Day Reconstruction Method (Kahneman, Krueger, Schkade, Schwarz, & Stone, 2004) exploits this process to collect emotion experience data. Because each day

contains many such episodes, the mind uses emotion as a switch to decide if an episode is important enough to respond to with a change in behavior. When emotional arousal during an episode is low, the episode does not strongly alter behavior and is soon forgotten. Above a certain threshold of emotional arousal, a change in behavior is motivated, and the episode is remembered. Thus, the remembered episodes over the course of an experience aggregate to form an overall evaluation. While this model explains the importance of emotion in overall evaluations, bridging the "experiencing self" with the "remembering self," it is unclear how the temporal dynamics of emotion affect how positive or negative the overall evaluation is.

An oft-cited theory attempting to explain overall evaluations based on emotion development over the course of an experience is peak-end theory. Peak-end theory posits that the best predictor of overall evaluation is an average of the strongest emotional moment, and the last emotional moment, of that experience (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993). This theory is largely based on studies of brief, simple experiences of pain, and does not hold for longer, more complex experiences such as workaday life or tourism, where the average of emotions experienced is a better predictor (Kemp, Burt, & Furneaux, 2008; Miron-Shatz, 2009; Strijbosch et al., 2019). These findings suggest that the more positive episodes that exist during an experience, regardless of their order, the more positive the overall evaluation will be. The cited findings are based on self-report, however, which has well-known measurement biases (Bastiaansen et al., 2019; Wirtz, Kruger, Scollon, & Diener, 2003). Therefore, Bastiaansen et al. (2019) recommend the use of passive physiological measurement to capture emotion as it develops through multiple episodes over time, a method which is starting to appear in tourism research (Kim & Fesenmaier, 2015; Shoval, Schvimer, & Tamir, 2018) and spatial experience research in geography (Birenboim, Dijst, Scheepers, Poelman, & Helbich, 2019).

The peak model of tourist emotions (Mitas, Yarnal, Adams, & Ram, 2012) suggests that tourists' *positive* emotions in particular develop according to a 'peak' pattern, wherein they are highest at the middle of the experience (Nawijn, Mitas, Lin, & Kerstetter, 2013). While the peak model explains how tourist experiences influence the course of emotions across multiple episodes and days, it does not explain how these emotions in turn influence overall evaluations. While a "peak" in positive emotions seems typical in tourism experiences, it is not known if this pattern is optimal.

Several cross-sectional studies show that recalled positive emotions are associated with positive overall evaluations (Hosany, Prayag, Deesilatham, Cauševic, & Odeh, 2015; Prayag et al., 2013), but do not reveal which temporal dynamics—peak or otherwise—of positive emotions add up to the most positive overall experience. The Breda model suggests that the frequency and level of positive emotions across episodes is crucial, so that simply more positive emotion is better. This assumption is untested in tourism experiences, however. Thus, it remains unknown whether the experience with the most and strongest positive emotional moments also produces the most positive overall evaluation. In other words, it is unclear how moments of positive emotion are "dosed" over the course of an experience at optimal levels of overall evaluation. In the present study, we use a biometric measurement of emotional engagement over the course of a multi-part tourism experience to assess the relationship between emotional engagement over time and space and overall evaluation.

Our study

We measured 25 participants' location, skin conductance, and overall experience evaluation at the Vincentre visitor center and guided outdoor tour in the village of Nuenen, the

Netherlands. The famous Dutch painter Vincent van Gogh lived in Nuenen during a crucial formative period in his life as an artist. Having lived in a remarkable variety of Northern European cities and towns during various abortive career attempts, van Gogh returned to live in Nuenen with his parents while trying his hand at painting. During his stay in Nuenen, he developed his distinctive painterly style and his focus on agricultural scenes and common people as subjects. The Vincentre does not display Van Gogh's original paintings but it features a number of interactive displays documenting Vincent's life as a young man and a painter at the beginning of his artistic career.

Research questions

Our main research question was: what course of emotion development over time and space makes for the best possible experience? To address this question, we formulated several specific research questions:

1) How were the visits to the Vincentre and Nuenen distributed over time and space?
2) To what extent did visitors to the Vincentre and Nuenen emotionally engage in the experience?
3) How did visitors evaluate the experience overall?
4) How did emotional engagement during the experience as a whole relate to overall evaluation?
5) Which exhibits during the experience had the strongest emotional engagement?
6) How did emotional engagement across various exhibits of the experience relate to overall evaluation? In other words, at which locations was emotional engagement the *most effective* in creating a positively evaluated experience?

Due to the minimal and inconsistent theoretical development around the unfolding of experiences over time and space, we took a hierarchical approach to answering our main question: what course of emotion development over time and space makes for the best possible experience? First, we examined the average between-participant experience and evaluation. Then, we modeled emotional engagement during the experience as a function of overall evaluation for each of the four two-dimensional spaces—the three floors of the museum, plus the outdoor tour in the village. To these models, we then added proximity to each specific exhibit to determine which were strongest in emotional engagement. Finally, we added interaction terms between overall evaluation and proximity to each location to determine at which exhibits emotional engagement was most *effective* in making the overall experience positive.

Methods

Sample

We used a convenience sample of 25 Master of Arts students in a leadership-themed study program based in the south of the Netherlands. The sample included a wide range of nationalities, including 13 from the Netherlands, 7 from other European countries, and 5 from South American and Asian origins. The sample was mostly female (21 of 25; 84%). Participants ranged in age from 21 to 51 years (mean = 30 years). According to conversations with Vincentre staff, this comprised a somewhat younger and more female sample than the overall visitor

population of the Vincentre, although the distribution of nationalities was reasonably representative.

Data collection

We asked participants to visit the Vincentre museum in three shifts of 8 participants each for a maximum of 30 minutes. During their museum visit, each participant was able to choose which exhibits they viewed and set their own pace and sequence to the visit. Historical objects, interpretive explanations, and digital interactive displays which participants could view are henceforth referred to as "(indoor) exhibits." Each shift then followed an outdoor guided tour of attractions related to Van Gogh in the village of Nuenen. Nominally also 30 minutes, these tours actually took 45 to 50 minutes per group and featured stories of Van Gogh's life in Nuenen. Stories addressed the locations he painted and how they have changed over the years, as well as Van Gogh's often difficult social and family relations. Each stop on the tour which included a story or explanation by the tour guide are henceforth referred to as "(outdoor) exhibits." At the beginning and end of the overall experience, including indoor and outdoor exhibits, participants filled out questionnaires about their experience evaluations. The pre-visit questionnaire also informed participants about the procedure and asked for their consent. Participants' first names were used to link their digital records (questionnaires, wearable, and location) with one another.

After filling out the intake questionnaire, participants were fitted with Empatica E4 wearable devices to measure their skin conductance. The E4 devices were fitted to participants' non-dominant wrist and fastened tightly. Indoor location tracking was engaged on smartphones which were then placed in participants' pockets. Finally, the button interface on the Empatica E4 was pressed to create a time marker indicating the beginning of the visit. Additional time markers were also created as each participant finished the indoor visit and began the outdoor tour. Between these two markers, indoor location tracking was terminated and outdoor location tracking started. After the outdoor tour, the beginning procedure was reversed: a fourth and final time marker was made, location tracking ended, smartphone and wearable collected back from participant, and final questionnaire administered.

Measures

Questionnaires at the beginning and end of the overall experience addressed participants' current emotional state, intent to recommend, and several overall evaluative terms. Over the course of the experience, we used wearable wristbands to record participants' skin conductance as a continuous, time-mapped measure of their state emotional engagement.

Emotional engagement. We used skin conductance responses as a measure of participants' state emotional engagement. Skin conductance responses refer to rapid, short-term increases in the ability of skin on the hands and feet to conduct electricity based on the opening of eccerine sweat glands in response to emotional stimuli (Mauss & Robinson, 2009). The raw level of skin conductance was measured using Empatica E4 wearable wristbands, which record the level of conductance between two electrodes on the bottom of the wrist at a frequency of 4 Hz. To derive skin conductance responses from this signal, we followed the procedures recommended by Benedek and Kaernbach (2010) and Braithwaite, Watson, Jones, and Rowe (2015). First, the data were imported into Matlab and cut according to recorded start and end times of each participant's indoor and outdoor visits. Next, data were visualized in Matlab and cleaned of artifacts using a custom Matlab script *ArtifactZ* which highlights any deviations greater than a threshhold (set in this case to 3 standard eviations) during a brief, moving time

window (in this case set to 20 seconds, based on trying longer and shorter time windows and finding fewer artifacts). Highlighted artifacts were visually inspected and when they appeared to be physiologically impossible (steep declines after a rise, which normally last at the very least 3 seconds) or coincided with large physical motions according to accelerometer data, were replaced by simple linear interpolation.

Then, skin conductance data were separated into slow-long term changes, known as the tonic signal or skin conductance level, and rapid short-term changes due to central nervous responses to emotional experiences, also known as the phasic signal or skin conductance responses (Benedek & Kaernbach, 2010). A mathematical procedure known as continuous deconvolution, implemented in the open source Matlab toolbox Ledalab, was used for this step. Further analyses involved the phasic signal, as this signal best reflects the response of skin conductance to emotional experience.

Location tracking. We tracked participants' indoor location using 25 Estimote Bluetooth beacons placed throughout the museum and Motorola Moto E5 (XT1944-2) smartphones with an application designed by one of the authors. The application recorded signal strength from all recognized beacons at a frequency of 1 Hz. Due to the multitude of walls and display cases, 25 beacons were marginally sufficient for the smartphones to consistently recognize signals from at least 3 beacons, a necessity for pinpointing their precise location. Thus, for visualizing the data, we created a script that estimated location based on the 3 nearest beacons, while for statistical analysis, we used only signal strength from the nearest beacon to reduce the amount of missing data.

We tracked outdoor location using signals from Global Positioning System (GPS) satellites using Strava, a popular smartphone application used for recording workout performance. As with indoor location data, recording frequency was 1 Hz. GPS coordinates were converted to latitude-longitude format for further analysis.

Overall experience evaluations. The quality of tourism experiences is often evaluated using a single item which measures intent to recommend. We measured this using a 0-10 scale (Reichheld, 2003). Using the same scale, we also asked for several other overall experience evaluations, asking to what extent participants *appreciated, enjoyed, found meaningful*, and *intended to revisit* the attractions. A Cronbach alpha of 0.93 among these 5 items suggested that they all measured a single underlying construct. Thus, we conducted the analyses using only the single item of intent to recommend, to enhance comparability with other applied and academic studies.

Data loss

It is important to note that our data collection suffered from significant data loss due to the myriad technologies used. Technical problems caused us to lose 7 participants due to missing indoor location data, 6 participants due to missing outdoor location data, and 2 (outdoor) to 5 (indoor) participants due to missing skin conductance data. Additionally, 2 participants were eliminated because, although their skin conductance signal appeared to be valid (non-zero), it did not show enough change over time to suggest the participant's sweat glands respond to emotions in experiences. This is a common limitation in skin conductivity research, and most general population estimates of such "non-responders" are around 10% (Braithwaite et al., 2015). Thus, despite starting with 25 participants, our models are based on 15 participants for the outdoor portion of the experience, and 12 participants for the indoor portion.

Although the number of participants was relatively low, high frequency of data collection resulted in high resolution spatial and temporal data for each participant, with over 200000 measurements acquired outdoors and over 35000 measurements collected indoors. Thus, we were still able to use over 100000 physiological measurements and over 30000 locations.

Analysis

To address research questions 1-3, we first described and visualized the spatial distribution of the data. Subsequent research questions implied examining the relationships among location, emotional engagement, and overall experience evaluation.

Experiences in general. We created kernel density maps of each 2-dimensional space using GRASS 7.6 to show where data points were most dense, reflecting where the largest number of participants spent the most time. We computed kernel density estimate maps of measurements acquired within each 2-dimensional space using the GRASS GIS 7.6 module v.kernel (Menegon and Blazek, 2019) with empirically selected Epanechnikov (parabolic) kernel function. Given the constant time interval of data acquisition, the maps of measurement densities reflect the cumulative time that the participants spent at any given location. This allowed us to identify locations where the largest number of participants spent the most time.

We computed maps of mean phasic skin conductance for each grid cell for each participant to understand the between participant variability in the measured values across the studied spaces. We then computed the mean for all participants at each grid cell: we used a 4 meter grid cell for outdoors and 1 a meter grid cell for indoors maps. To visualize the distribution of mean skin conductance we converted the gridded means to points and interpolated high resolution, smooth representation of the means using regularized spline with tension method implemented in GRASS GIS 7.6 module v.surf.rst (Mitas et al., 2019). Then, to address research questions 2-3, we then conducted appropriate descriptive statistics for each variable of interest.

Modeling proximity and recommendation effects on emotion. Mixed-effects linear models were used with emotional engagement as a continuous outcome variable. While it is more consistent with established theory (Bastiaansen et al., 2019) to think of emotional engagement as mediating the effects of a particular exhibit on experience evaluation, the small sample size made it difficult to model overall evaluation, a between-participants variable, as the outcome variable. Thus, we used emotional engagement, which varied in time and space as well as between participants, as the outcome variable. As the models included both between- and within-participant effects, used a random intercept to control for different baseline levels of emotion engagement between participants.

We built models hierarchically. Initial models addressed the between-participant relationship between overall evaluation and emotional engagement. Subsequently, predictors based on proximity to each exhibit were added to each model. Proximity to each exhibit was calculated according to a Gaussian curve and ranged between 1, in which case participants were standing directly at the exhibit, and approaching 0, in which case participants were more than 30 meters away from the exhibit (at the outdoor tour) or more than 6 meters away (inside the museum). These models allowed us to determine which locations were most emotionally engaging.

The final step in our analysis was to add first-order interaction terms between each proximity predictor and intent to recommend. These predictors essentially show if the changes in emotional engagement as participants neared an exhibit differed for participants of different levels of intent to recommend. In other words, if the interaction between intent to recommend

and proximity to a certain exhibit was positive, becoming more emotionally engaged near this exhibit was related to a positive evaluation. We term this *emotional effectiveness*, as it reflects the extent to which emotions at a particular exhibit were part of a more positive overall experience.

To visualize the modeled effects, we first computed the Pearson product-moment correlation between each participant's average skin conductance within a grid cell and their intent to recommend. We then created high resolution smooth representation maps of the correlation similarly as for the mean skin conductance by interpolating the low resolution gridded values to higher resolution. Positive correlations were shaded red, while negative correlations were shaded blue, with saturation of color indicating the magnitude of the correlation. Thus, squares with correlation near 0 displayed white.

Results

Experiences in general

Our first research question asked where participants spent their time. The density map of the outdoor data (Figure 1) reflects the eight stops on the tour, with eight strong concentrations of points showing on the maps: the Vincentre museum, where tours started and ended; the vicarage, where Van Gogh's father and mother lived; the path behind the vicarage, showing the shack where Van Gogh stayed during his time in Nuenen; the pond behind this area, which Van Gogh painted; the former post office, where Van Gogh exchanged letters with his brother; the church which Van Gogh also famously painted; an old tree on the central village square; and finally, a mansion called Nune Ville where a love interest of Van Gogh lived. The peaks of these eight concentrations of data points were extracted used as locations of outdoor exhibits for further analysis, as they indicate where participants stood to listen to the tour guide present each specific exhibit.

The density maps of the indoor data (Figure 2) show that most data points occurred in the main areas of passage, that is, stairwells and the central part of each floor, rather than very near the exhibits. This may also be due to the imprecise estimation of the location of each data point, as glass and concrete surfaces in the museum somewhat reduced the strength of beacon signals.

FIGURES 1 and 2 ABOUT HERE

Emotional engagement varied across the four 2-dimensional spaces of the experience. It was highest outdoors (mean = 0.053, sd = 0.06), followed by the ground floor of the museum (mean = 0.048, sd = 0.07), the first floor (mean = 0.039, sd = 0.04), and the second floor (mean = 0.034, sd = 0.04). The relatively large standard deviations reflect the nature of phasic skin conductance, which has a baseline near 0 but spikes at moments of emotional arousal, creating a relatively low mean with high variation. It is worth noting that the floors of the museum were typically visited from the ground floor upwards, so the museum visit appears to have become less emotionally engaging over time, to be later surpassed by the outdoor visit (Figures 3 and 4).

FIGURES 3 and 4 ABOUT HERE

Participants' overall evaluations of the experience spanned the full 11-point range of intent to recommend, from 0 to 10. The most frequent responses, with three participants each,

were 5, 6, and 9. The mean was 6.13 with a standard deviation of 2.8. As a general evaluation of a tourist attraction in the Netherlands, this outcome would be considered rather low, as high-performing tourist attractions usually report intent to recommend or other overall evaluation scores around 8 to 8.5.

Proximity and recommendation effects on emotion

      Recommendation and emotion. We initially modeled emotional engagement as a function of intent to recommend, a purely between-participants analysis, for each two-dimensional space (Table 1). Not surprisingly, none of these relationships approached significance, as the sample sizes of 12 and 15 participants afforded insufficient statistical power. It is important to note, however, that the signs of the coefficients differed across the four spaces. Zero-order Pearson product-moment correlations (r) between the average phasic level per participant and intent to recommend are also given to illustrate the magnitude of these effects. Intent to recommend had negative coefficients on emotional engagement outdoors ($r = -0.11$) and on the ground floor ($r = -0.24$), and positive coefficients on the first ($r = 0.13$) and second ($r = 0.43$) floors. In other words, participants who evaluated their overall experience most positively were not more emotional throughout. Rather, they were less emotional at the beginning (ground floor of the museum) and end (outdoor tour), and more emotional in the middle (first and second floors of the museum). A larger sample would be needed to determine if these effects are statistically significant.

<div align="center">TABLE 1 ABOUT HERE</div>

      Proximity and emotion. We subsequently added predictors based on proximity to each exhibit to the models. The resulting models each significantly ($p < 0.0001$) fit the data better than a model with only intent to recommend as a predictor (Table 2). These models demonstrate the changes in emotional engagement with increasing proximity to each exhibit within the average participant, while controlling for the differing baselines in emotional engagement between participants. Exhibit locations corresponded to locations of Bluetooth proximity beacons at the museum, and to the eight highly dense concentrations of data points during the outdoor tour. On the ground floor, proximity to an interactive exhibit showing a map on a pull-down curtain ($p < 0.05$) and to the right side of the Faces exhibit ($p < 0.0001$) were relatively more emotionally engaging. The Faces exhibit shows computer-animated period portraits of faces from Van Gogh's life in Nuenen. In contrast, the left side of the Faces exhibit and the entry were relatively less emotionally engaging. There was no significant effect of being relatively nearer the exit (Table X). It is worth noting that, as the map (Figure 6) shows, the entry and exit are actually opposite sides of a single portal, and thus rather near each other.

      On the first floor, a passage near the loom ($p < 0.0001$), the door to the non-functioning movie theater ($p < 0.0001$), a corner with the outside wall of the building ($p < 0.0001$), cupboards showing period objects such as medicines and paints ($p < 0.01$), and a display of period hats ($p < 0.0001$) were relatively less emotionally engaging. The corner between the hats and the cupboards ($p < 0.0001$), a window toward the street ($p < 0.0001$), and the hall connecting most of the major exhibits ($p < 0.05$) were relatively more emotionally engaging. Proximity to an interactive shadowbox exhibit and to the loom had no significant effect on emotional engagement. These were incidentally the first two exhibits participants on the first floor saw.

      On the second floor, the stairs between the first and second floors and a table with an interactive drawing activity were less emotionally engaging, while an exhibit with videos of Van

Gogh's present-day family relatives was relatively more emotionally engaging (all p's < 0.0001). Finally, during the outdoor tour of the village, six of the eight exhibits, namely the vicarage, path, pond, post office, tree, and Nune Ville were relatively less emotionally engaging (all p's < 0.0001). Proximity to the church and the Vincentre had no significant effect on emotional engagement.

TABLE 2 ABOUT HERE

Emotional effectiveness. The final step in our modeling process was to add first-order interaction terms between proximity to each exhibit and the intent to recommend the overall experience. These predictors reflect *emotional effectiveness*, the extent to which emotional engagement near a particular exhibit was part of a positive overall experience (Table 3). Once again following the chronology of participants' visits, we first discuss models of the indoor part of the experience, starting with the ground floor.

On the ground floor, only the entry to the museum exhibits was positively effective, while the exit, both sides of the Faces exhibit, and the pull-down map were negatively effective (all p's < 0.0001 except Faces right, p < 0.01). Thus, highly recommending visitors became more emotional near the entry and less emotional near the other location beacons on the ground floor (Figure 14). On the first floor (Figure 15), all exhibits except the cupboards exhibit featured significant interactions with intent to recommend. The emotional effectiveness represented by these interactions was positive for the back of the loom (p < 0.05), the window to the street (p <0.0001), the corner between the cupboards and the hats (p < 0.01), and both sides of the shadowbox exhibit (p < 0.0001 and p < 0.05). Emotional effectiveness was negative for the passage near the loom (p < 0.05), the defunct theater (p < 0.0001), the corner between the cupboards and outside wall (p < 0.0001), and the hats exhibit (p < 0.0001). Moving on to the second floor (Figure 16), all four beacon locations featured significant emotion effectiveness, with the west side of the family exhibit being positively effective and the east side of the family exhibit, stairwell, and drawing table being negatively effective (all p's < 0.0001 except family east, p < 0.05). In other words, the latter three exhibits were less emotional for participants who reported higher intent to recommend, while the west side of the family exhibit was more emotional for those participants (Figure 6).

During the outdoor tour (Figure 5), all exhibits except for the post office had significantly negative emotional effectiveness (all p's < 0.0001 except the pond, p < 0.01). The post office was positively effective (p < 0.05). Compared to the indoor models, exhibits were much further apart, so the proximity predictors must be interpreted relative not only to each other, but relative to the locations recorded *between* exhibits, that is, while participants were walking from one exhibit to the next. Thus, more highly recommending participants felt not only less emotional at the other seven exhibits compared to the post office, but also compared to the spaces of the tour between exhibits.

TABLE 3 ABOUT HERE
FIGURES 5 and 6 ABOUT HERE

Discussion

We recorded skin conductance and intent to recommend from 25 visitors to the Vincentre museum and Vincent van Gogh-themed outdoor village tour in Nuenen, the Netherlands. Of these, 15 provided usable data, comprising over 100,000 measurements. We modeled their emotional engagement, operationalized as phasic skin conductance, as a function of intent to recommend. We then added proximity to various exhibits, and finally the interactions between intent to recommend and proximity. We termed the latter interactions as representing the *emotional effectiveness* of each exhibit. Our findings show that participants found some exhibits, and some two-dimensional spaces—the floors of the museum and the outdoor portion—as substantially more emotionally engaging and more emotionally effective than others. In fact, two of the four spaces and 17 out of 28 total exhibits had negative effectiveness, meaning that participants who became less emotionally engaged nearby had an experience they would more likely recommend. In short, the best experiences were not always the most emotional. More precisely, the best experiences were more emotional in certain locations and not in others.

The wide range of intent to recommend in the sample suggests that not all of the 15 analyzed participants had positive experiences, although some clearly did. The pattern of emotional engagement across the four spaces is the opposite of the peak model of positive emotions (Mitas, Yarnal, Adams, et al., 2012). Several studies have shown that positive emotions are elevated during a tourism experience compared to before and after, and indeed also elevated during the middle part of a tourism experience compared to the beginning and end (Nawijn et al., 2013). There could be a number of reasons for this discrepancy. Nawijn and Fricke (2015) have challenged the applicability of the peak model to dark tourism experiences, which focus largely on death and suffering. Because of the social ostracism Van Gogh faced during his Nuenen residency, some participants may have experienced their visit as dark tourism, or in any case something 'darker' than the experiences the peak model was built on. A second possibility is that the peak model, which even in the briefest previous studies was assessed over a multi-day span of time, does not hold for an experience totaling less than two hours.

It is more likely, however, that the emotional engagement as we measured it—according to phasic skin conductance, which spikes whenever an individual feels *any* emotional arousal, positive or negative—reflected a mix of positive and negative emotions. It is quite possible that the *positive* emotions participants felt during their visit did follow the peak model. Consistent with this possibility, while emotional engagement followed the opposite of the peak model, emotional effectiveness followed the peak model, with effectiveness being negative in the first and last spaces participants visited, and positive in the middle two spaces.

In other words, the emotional engagement of the higher-recommending participants did follow the peak pattern. This could be because the emotional engagement of higher-recommending participants reflected primarily positive emotions, which would be consistent with the substantial literature linking recommendation to positive emotions (e.g., Prayag et al., 2013). It is also possible that, regardless of emotion valence, higher recommending participants were more strongly moved by the exhibits of the first and second floors than of the ground floor or the outdoor tour. A pattern in the specific exhibits which had positive or, conversely, negative effectiveness could not be discerned. Future qualitative research allowing participants to retell their experiences while judging the various exhibits could reveal such patterns.

Based on these findings, we assert that the best experience is not simply the most emotional experience, and that while a certain threshold of emotion drives experience outcomes such as recommendation (Prayag et al., 2013), more emotion is not necessarily better. Not only between the floors of the museum and the outdoor tour, but also within each of these spaces,

there were substantial differences in the emotional effectiveness of various exhibits. The most positively evaluated experiences clearly comprised a mix of more and less emotionally intense moments. Thus, we believe the model of experience presented in Bastiaansen et al. (2019) should be revised. Specifically, the model should reflect that the influence of emotion on overall evaluations is not linear, and depends to some extent on the emotion experienced and, especially, on the dynamics of how various moments during an experience trigger and attenuate emotion over time. Our data show that in this process some moments should be emotional, while others *should not*, to combine into an optimally evaluated experience.

Limitations

The conclusions and contributions of our research carry several limitations. The limitation which introduced the most ambiguity into the findings is the lack of a clear metric of emotional valence from the physiological signals we have interpreted. We focused on skin conductance as a measure of emotional engagement. Skin conductance does not, however, distinguish between positive and negative emotion. There are some hints that heart rate, temperature, or heart rate variability might, but understanding of these relationships is extremely tentative (Kreibig, 2010).

We were also limited by a non-generalizable convenience sample and small sample size which was made smaller by substantial data loss. Thus, our power to detect between-participant differences, which were necessary to assess overall experience evaluations, was limited. A larger sample which could be demographically matched to a population of interest would make future studies of this sort more robust and generalizable.

Conclusion

Our study has confirmed the important role of emotion in understanding why some peoples' tourism experiences turn out better than others'. Furthermore, we have demonstrated that despite suggestion in previous literature that emotion is good in tourism experiences, more emotion is not necessarily better. Indeed, the best possible experience seems to be composed from a mix of more- and less-emotionally engaging elements, rather than one continuous emotional high.

While our data shed light on the specific dynamics that make a visit to Nuenen as positive as possible, the extent to which these dynamics are generalizable to other experiences is an issue for future research. There is not enough evidence to conclude, for example, that the best experiences tend to be emotional in the middle, although this would be somewhat consistent with research on the peak model. Our data also do not show which specific emotions, or what sequence of positive, negative, and unemotional moments, makes an experience optimally positive. To support such conclusions, methods such as those we used in Nuenen, plus a measure of valence by self-report, facial expression measurement, or as yet unknown physiological marker of valence would be needed.

References

Bastiaansen, M., Lub, X., Mitas, O., Jung, T. H., Passos Acenção, M., Han, D., . . . Strijbosch, W. (2019). Emotions as core building blocks of an experience. *International Journal of Contemporary Hospitality Management, 31*.

Benedek, M., & Kaernbach, C. (2010). Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology, 47*(4), 647-658.

Birenboim, A., Dijst, M., Scheepers, F. E., Poelman, M. P., & Helbich, M. J. T. P. G. (2019). Wearables and location tracking technologies for mental-state sensing in outdoor environments. 1-13.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2015). *A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments*. Retrieved from Birmingham, UK:

Hosany, S., & Gilbert, D. (2010). Measuring tourists' emotional experiences toward hedonic holiday destinations. *Journal of Travel Research, 49*(4), 513-526.

Hosany, S., Prayag, G., Deesilatham, S., Causevic, S., & Odeh, K. (2015). Measuring tourists' emotional experiences: Further validation of the destination emotion scale. *Journal of Travel Research, 54*(4), 482-495.

Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. J. P. s. (1993). When more pain is preferred to less: Adding a better end. *4*(6), 401-405.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science, 306*(5702), 1776-1780.

Kemp, S., Burt, C. D., & Furneaux, L. (2008). A test of the peak-end rule with extended autobiographical events. *Memory & Cognition, 36*(1), 132-138.

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

Knobloch, U., Robertson, K., & Aitken, R. (2017). Experience, Emotion, and Eudaimonia: A Consideration of Tourist Experiences and Well-being. *Journal of Travel Research, 56*(5), 651-662.

Kreibig, S. D. J. B. p. (2010). Autonomic nervous system activity in emotion: A review. *84*(3), 394-421.

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.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & emotion, 23*(2), 209-237.

Miron-Shatz, T. J. E. (2009). Evaluating multiepisode events: Boundary conditions for the peak-end rule. *9*(2), 206.

Mitas, O., Yarnal, C., Adams, R., & Ram, N. (2012). Taking a "peak" at leisure travelers' positive emotions. *Leisure Sciences, 34*(2), 115-135.

Mitas, O., Yarnal, C., & Chick, G. (2012). Jokes build community: Mature tourists' positive emotions. *Annals of tourism research, 39*(4), 1884-1905.

Moyle, B. D., Moyle, C.-l., Bec, A., & Scott, N. (2017). The next frontier in tourism emotion research. *Current Issues in Tourism*, 1-7.

Nawijn, J., & Fricke, M. C. (2015). Visitor emotions and behavioral intentions: The case of concentration camp memorial Neuengamme. *International Journal of Tourism Research, 17*(3), 221-228.

Nawijn, J., Mitas, O., Lin, Y., & Kerstetter, D. (2013). How do we feel on vacation? A closer look at how emotions change over the course of a trip. *Journal of Travel Research, 52*(2), 265-274.

Pearce, P. L., & Zare, S. (2017). The orchestra model as the basis for teaching tourism experience design. *Journal of Hospitality and Tourism Management, 30*, 55-64.

Prayag, G., Hosany, S., & Odeh, K. (2013). The role of tourists' emotional experiences and satisfaction in understanding behavioral intentions. *Journal of Destination Marketing & Management, 2*(2), 118-127.

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

Scott, N., & Le, D. (2017). Tourism experience: a review. In N. Scott, J. Gao, & J. Ma (Eds.), *Visitor Experience Design* (Vol. 5, pp. 30). Wallingford, Oxfordshire ; Boston, MA: CABI.

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.

Skavronskaya, L., Scott, N., Moyle, B., Le, D., Hadinejad, A., Zhang, R., . . . Shakeela, A. (2017). Cognitive psychology and tourism research: state of the art. *Tourism review, 72*(2), 221-237.

Strijbosch, W., Mitas, O., van Gisbergen, M., Doicaru, M., Gelissen, J., & Bastiaansen, M. J. F. i. p. (2019). From experience to memory: on the robustness of the peak-and-end-rule for complex, heterogeneous experiences. *10*.

Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science, 14*(5), 520-524.

Table 1. Correlation coefficients between intent to recommend and emotional engagement

Space	Pearson's r	p
Ground floor	-0.24	0.418
First floor	0.13	0.661
Second floor	0.43	0.143
Outdoor	-0.11	0.686

Table 2. Models of emotion engagement as a function of proximity

Space	Predictor	Fixed effect coefficient estimate	SE	T	Model AIC
Ground floor					-49491.9
	(Intercept)	0.11283	0.04872	2.32*	
	Intent to recommend	-0.00308	0.00685	-0.45	
	Entry	-0.03980	0.00357	-11.14***	
	Exit	-0.00564	0.00360	-1.57	
	Faces exhibit	-0.06682	0.00358	-18.67***	
	Right side of Faces	0.01855	0.00374	4.96***	
	Pull down exhibit	0.00741	0.00294	2.53*	
First floor					-74624.39
	(Intercept)	0.06474	0.02899	2.23*	
	Intent to recommend	0.00279	0.00418	0.67	
	Loom	-0.00260	0.00210	-1.24	
	Hall next to loom	-0.01479	0.00348	-4.25***	
	Theater	-0.02457	0.00354	-6.94***	
	Window	0.01550	0.00370	4.19***	
	Corner near window	-0.04730	0.00416	-11.37***	
	Cupboards	-0.01148	0.00420	-2.73**	
	Hat exhibit	-0.03604	0.00384	-9.38***	
	Corner of hats and cupboards	0.03063	0.00292	10.49***	
	Shadowboxes	-0.00545	0.00327	-1.67.	
	Hall next to shadowboxes	-0.00121	0.00216	-0.56	
	Center of hall	0.00660	0.00297	2.23*	

Space	Predictor	Fixed effect coefficient estimate	SE	T	Model AIC
Second floor					-59181.97
	(Intercept)	0.02099	0.02240	0.94	
	Intent to recommend	0.00431	0.00322	1.34	
	Stairs	-0.04459	0.00388	-11.49***	
	Family 1	0.06981	0.00350	19.94***	
	Family 2	0.03009	0.00401	7.51***	
	Drawing table	-0.10022	0.00374	-26.79***	
Outdoor					-397899.9
	(Intercept)	0.04480	0.01850	2.42*	
	Intent to recommend	-0.00102	0.00276	-0.37	
	Church	-0.00183	0.00101	-1.81.	
	Vincentre	0.00035	0.00104	0.34	
	Vicarage	-0.03685	0.00118	-31.28***	
	Pond	-0.02756	0.00076	-36.04***	
	Nune Ville	-0.03275	0.00128	-25.49***	
	Tree	-0.02798	0.00101	-27.70***	
	Path	-0.02556	0.00117	-21.92***	
	Post office	-0.02147	0.00084	-25.66***	

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Outcome variable: Emotional engagement measured by phasic skin conductance.

Table 3. Models of emotion engagement as a function of proximity x intent to recommend interactions

Space	Predictor	Coefficient estimate	SE	T	Model AIC
Ground floor					-49874.62
	(Intercept)	0.02714	0.04360	0.62	
	Intent to recommend (ITR)	0.00898	0.00614	1.46	
	Entry	-0.08585	0.01012	-8.48***	
	Exit	0.06215	0.01105	5.62***	
	Faces exhibit	0.07629	0.01569	4.86***	
	Right side of Faces	0.04893	0.01428	3.43***	
	Pull down exhibit	0.07378	0.01147	6.43***	
	Entry x ITR	0.00756	0.00142	5.34***	
	Exit x ITR	-0.01068	0.00157	-6.81***	
	Faces exhibit x ITR	-0.01747	0.00200	-8.74***	
	Right side of Faces x ITR	-0.00498	0.00193	-2.59**	
	Pull down exhibit x ITR	-0.00942	0.00145	-6.49***	
First floor					-75292
	(Intercept)	0.03599	0.02927	1.23	
	Intent to recommend (ITR)	0.00673	0.00419	1.60	
	Loom	-0.02471	0.01041	-2.37**	
	Hall next to loom	0.01579	0.01287	1.23	
	Theater	0.03330	0.01261	2.64**	
	Window	-0.08895	0.01262	-7.05***	
	Corner near window	0.09805	0.01337	7.33***	
	Cupboards	-0.00911	0.01403	-0.65	
	Hat exhibit	0.01747	0.01273	1.37	
	Corner of hats and cupboards	0.00236	0.01061	0.22	
	Shadowboxes	-0.08017	0.01193	-6.72***	
	Hall next to shadowboxes	-0.02268	0.00907	-2.50*	
	Center of hall	-0.00191	0.01003	-0.19	
	Loom x ITR	0.00293	0.00134	2.19*	
	Hall next to loom x ITR	-0.00383	0.00169	-2.26*	
	Theater x ITR	-0.00782	0.00165	-4.74***	
	Window x ITR	0.01341	0.00159	8.45***	
	Corner near window x ITR	-0.01947	0.00172	-11.35***	
	Cupboards x ITR	0.00067	0.00182	0.37	
	Hat exhibit x ITR	-0.00757	0.00168	-4.51***	
	Corner of hats and cupboards x ITR	0.00386	0.00141	2.74**	
	Shadowboxes x ITR	0.00980	0.00159	6.18***	
	Hall next to shadowboxes x ITR	0.00294	0.00118	2.49*	
	Center of hall x ITR	0.00082	0.00135	0.61	

Space	Predictor	Fixed effect coefficient estimate	SE	T	Model AIC
Second floor					-61687.79
	(Intercept)	-0.05012	0.01592	-3.15**	
	Intent to recommend (ITR)	0.01351	0.00227	5.95***	
	Stair	0.17061	0.01063	16.04***	
	Family 1	-0.27203	0.01093	-24.89***	
	Family 2	0.05987	0.01209	4.95***	
	Drawing table	0.20877	0.01089	19.16***	
	Stair x ITR	-0.02961	0.00134	-22.06***	
	Family 1 x ITR	0.04519	0.00142	31.93***	
	Family 2 x ITR	-0.00368	0.00157	-2.35*	
	Drawing table x ITR	-0.04226	0.00136	-31.09***	
Outdoor					-399001.5
	(Intercept)	0.03235	0.01891	1.71.	
	Intent to recommend (ITR)	0.00098	0.00282	0.35	
	Church	0.01184	0.00229	5.18***	
	Vincentre	0.02916	0.00204	14.30***	
	Vicarage	0.03192	0.00307	10.40***	
	Pond	-0.02188	0.00188	-11.66***	
	Nune Ville	0.01791	0.00365	4.91***	
	Tree	0.01350	0.00292	4.63***	
	Path	0.03046	0.00395	7.72***	
	Post office	-0.02472	0.00200	-12.34***	
	Church x ITR	-0.00215	0.00033	-6.54***	
	Vincentre x ITR	-0.00502	0.00033	-15.40***	
	Vicarage x ITR	-0.01044	0.00043	-24.23***	
	Pond x ITR	-0.00088	0.00029	-3.02**	
	Nune Ville x ITR	-0.00761	0.00051	-14.96***	
	Tree x ITR	-0.00626	0.00041	-15.28***	
	Path x ITR	-0.00864	0.00058	-14.81***	
	Post office x ITR	0.00064	0.00031	2.08*	

Notes: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1. Outcome variable: Emotional engagement measured by phasic skin conductance.

Figure 1. Density of data points during outdoor tour of Nuenen.

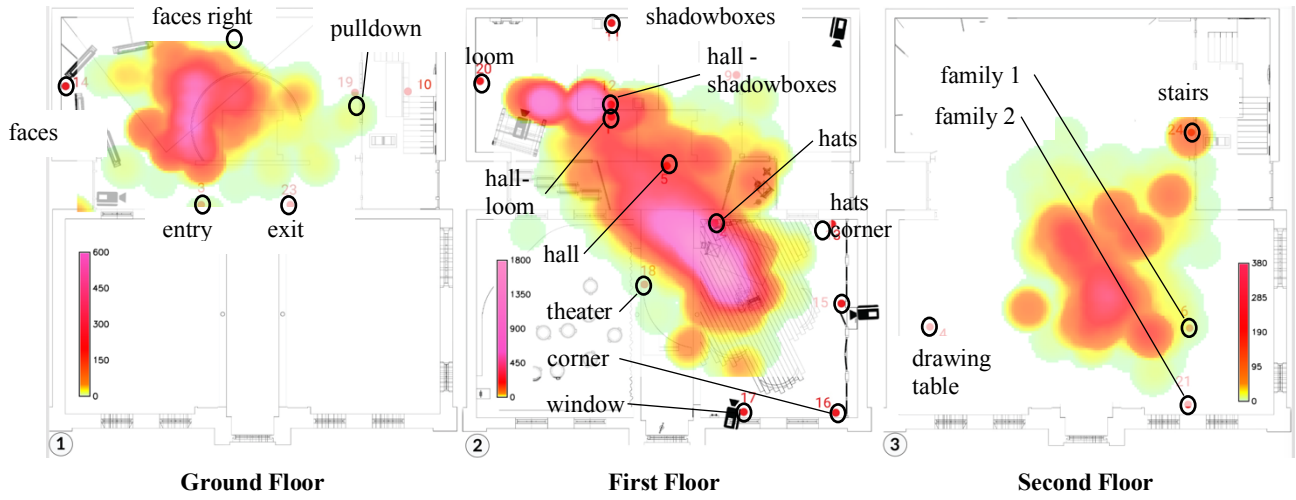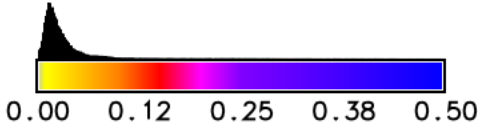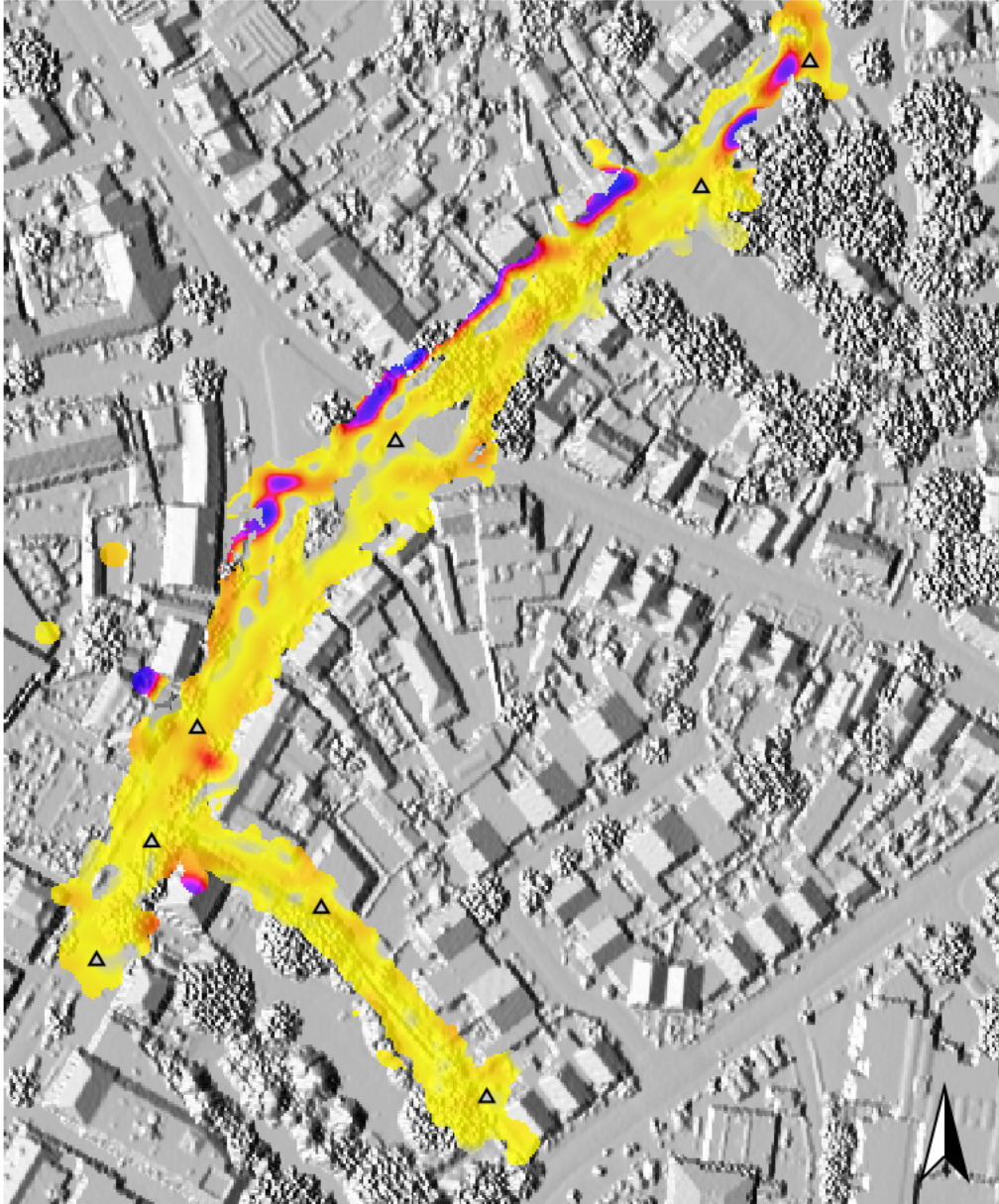Figure 2. Density of data points during the indoor visit of the Vincentre.

Ground Floor

First Floor

Second Floor

Figure 3. Outdoor phasic skin conductance levels.

Figure 4. Indoor phasic skin conductance levels.

Ground Floor   **First Floor**   **Second Floor**
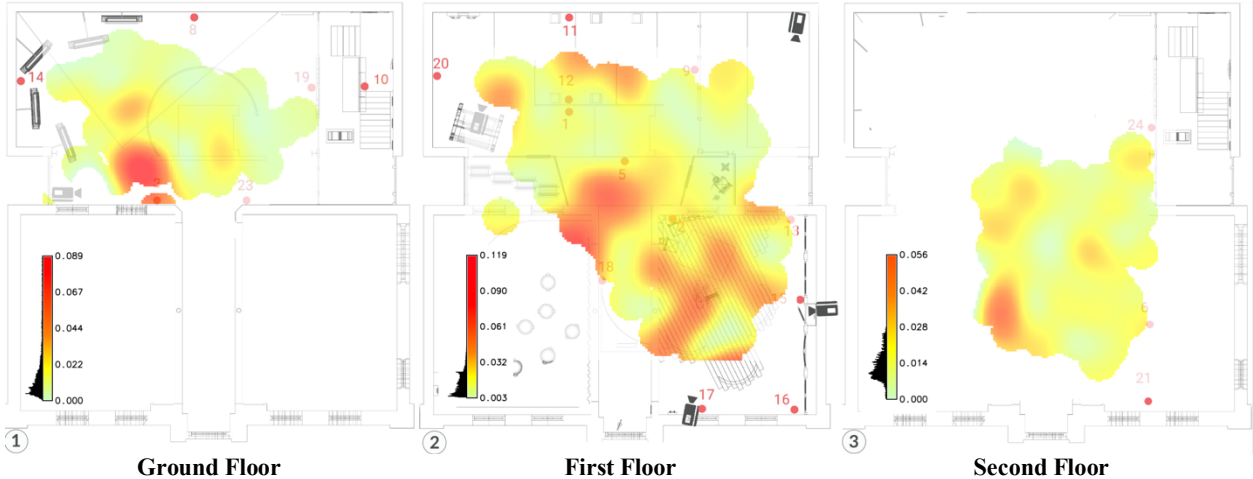
Figure 5. Correlations between phasic skin conductance and intent to recommend outdoors.
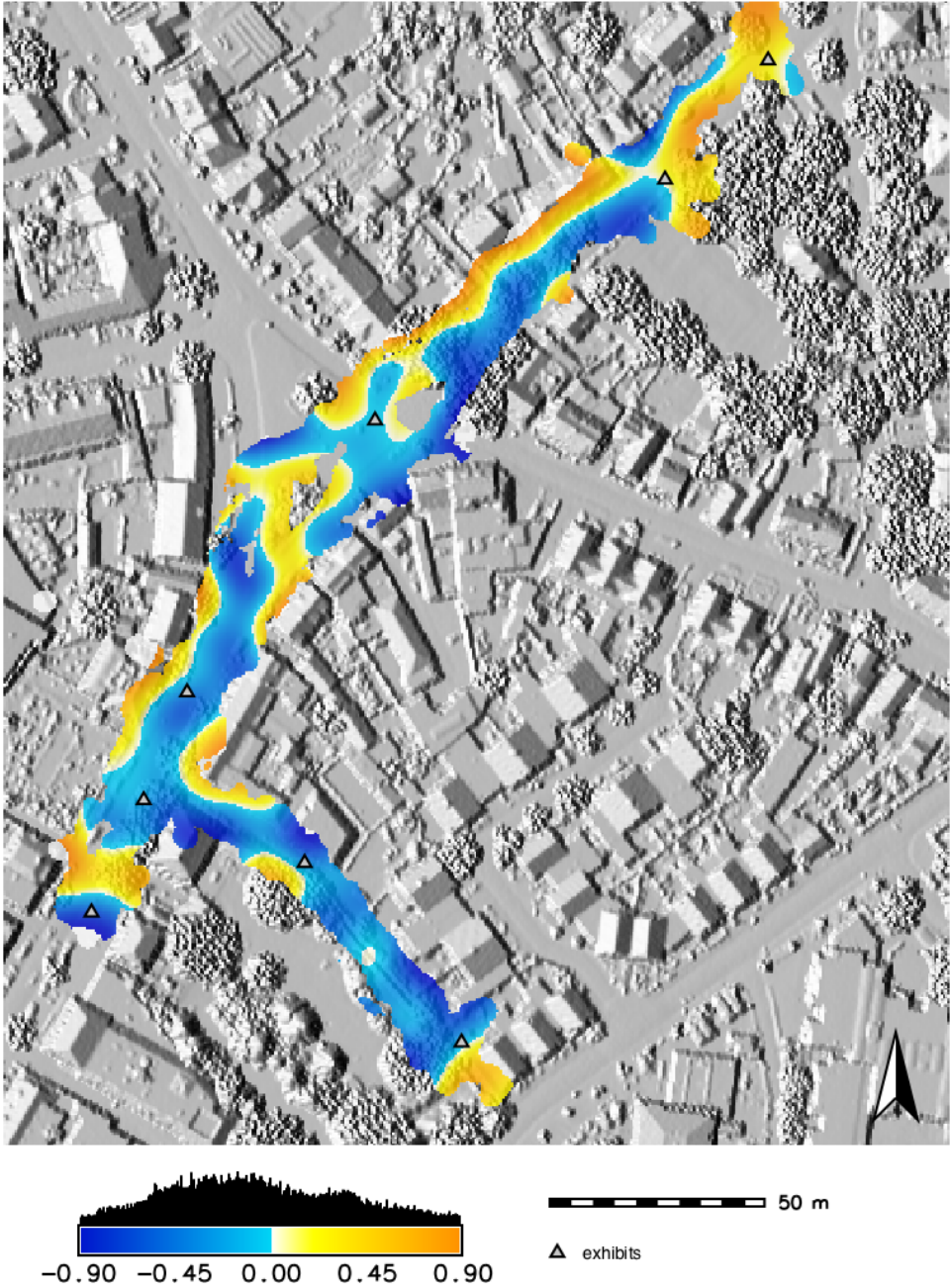
Figure 6. Correlations between phasic skin conductance and intent to recommend indoors.

Ground Floor   First Floor   Second Floor