Quantifying bicycling stress level using virtual reality and electrodermal activity sensor

Working Paper

Author(s):
Nazemi, Mohsen; van Eggermond, Michael A.B.

Publication date:
2020-08

Permanent link:
https://doi.org/10.3929/ethz-b-000446909

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 1568
**TRB Annual Meeting**

**Quantifying Bicycling Stress Level Using Virtual Reality and Electrodermal Activity Sensor**

--Manuscript Draft--

<table>
<thead>
<tr>
<th>Full Title:</th>
<th>Quantifying Bicycling Stress Level Using Virtual Reality and Electrodermal Activity Sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract:</strong></td>
<td>The application of wearable wireless electrodermal activity (EDA) devices in virtual reality (VR) experiments has become increasingly popular and lends itself for the application in behavioral research for transport planning. The type of infrastructure and interaction with other road users can invoke different arousal levels in urban bicyclists, which can be modeled with VR applications and captured by EDA sensors. At the intersection of engineering, psychology, and physiology, this research attempted to quantify bicycling stress levels using a bicycle simulator combined with immersive 360-degree virtual reality and an EDA sensor. Overall, 150 participants rode through 5 different bicycling environments in VR while their elicited skin conductance responses (SCRs) were passively collected by an EDA sensor that was connected to the participants for the duration of the experiment. Analysis of the signal for the entire stretch of the bicycling course did not yield significant differences in SCRs between different bicycling environments. However, comparing smaller segments of the bicycling course revealed significant differences. Bicycling on the sidewalk shared with pedestrians caused higher stress levels while bicycling on the segregated bicycle path found to be the least stressful. Evidence was found of a link between self-reported perceptions of safety and SCR rates. The results of this research shows a promising path in using VR experiments to identify stressful events and locations and to quantify bicycling stress level for non-existent future facility designs.</td>
</tr>
</tbody>
</table>

| Manuscript Classifications: | Planning and Analysis; Travel Survey Methods AEP25 |
| Manuscript Number: | |
| **Article Type:** | Presentation |
| **Order of Authors:** | Mohsen Nazemi, PhD Candidate |
| | Michael van Eggermond, PhD |
Quantifying Bicycling Stress Level Using Virtual Reality and Electrodermal Activity Sensor

This paper is being submitted in response to the call for papers - AR/VR and psychometric instruments in travel behavior research by the committee Travel Survey Methods (AEP25) (formerly ABJ40)

Mohsen Nazemi (Corresponding author)
Future Cities Laboratory
Singapore-ETH Centre
Singapore 138602
Singapore
nazemi@ivt.baug.ethz.ch

Michael A.B. van Eggermond
University of Applied Sciences Northwestern Switzerland (FHNW)
Hofackerstrasse 32, 4032 Muttenz
Switzerland
michael.vaneggermond@fhnw.ch

Word Count: 6216 words + 2 table(s) × 250 = 6716 words

Submission Date: August 1, 2020
ABSTRACT

The application of wearable wireless electrodermal activity (EDA) devices in virtual reality (VR) experiments has become increasingly popular and lends itself for the application in behavioral research for transport planning. The type of infrastructure and interaction with other road users can invoke different arousal levels in urban bicyclists, which can be modeled with VR applications and captured by EDA sensors. At the intersection of engineering, psychology, and physiology, this research attempted to quantify bicycling stress levels using a bicycle simulator combined with immersive 360-degree virtual reality and an EDA sensor. Overall, 150 participants rode through 5 different bicycling environments in VR while their elicited skin conductance responses (SCRs) were passively collected by an EDA sensor that was connected to the participants for the duration of the experiment. Analysis of the signal for the entire stretch of the bicycling course did not yield significant differences in SCRs between different bicycling environments. However, comparing smaller segments of the bicycling course revealed significant differences. Bicycling on the sidewalk shared with pedestrians caused higher stress levels while bicycling on the segregated bicycle path found to be the least stressful. Evidence was found of a link between self-reported perceptions of safety and SCR rates. The results of this research shows a promising path in using VR experiments to identify stressful events and locations and to quantify bicycling stress level for non-existent future facility designs.

INTRODUCTION

Bicycling risk and safety concerns have constantly been cited as one of the main barriers to bicycling. A lack of safety has kept bicycling a less attractive mode of transport in many places around the world. Understanding people’s perceptions of bicycling risk and assessing different bicycling environments quantitatively constitute a valuable input for transport and urban planning.

Bicycling risk can be quantified objectively by analyzing the number of crashes and incidents, (e.g. actual risk). It can also be studied subjectively by conducting different types of surveys and experiments, thus determining the perceived risk. However, scarcity of bicycling experience and bicycle infrastructure in cities with emerging bicycling cultures make objective assessment of road safety difficult, if not impossible. Furthermore, application of classic survey methods—including figures, images, or videos to describe different street designs—which heavily rely on respondents’ imagination of subjective bicycling risk, might not lead to adequate and reliable results. Therefore, it is necessary to overhaul the existing research methods in such conditions.

The term “stress” is commonly used in bicycle safety studies. It refers to the stress while bicycling, caused by passing motorized traffic or conflicts with pedestrians. Stress can be defined as “the non-specific mix of physiological and psychological responses of the body to any demand of change” (1). It is a reaction from a calm state to an aroused and alert state for the purpose of preserving the integrity of an organism and protecting oneself during a situation in which the subject feels threatened or attacked due to the presence of an stimulus. Experiencing stress can cause physiological reactions such as changes in heart rate (HR), blood pressure, breathing rate, and electrodermal activity (EDA) that can be measured by bio-sensors (2).

Bio-sensors have been used in bicycling research. HR was measured to study bicyclists’ perceived risk (3) and the influence of built environment on bicycling comfort (4). Eye-tracking was used to investigate bicyclists’ gaze behavior and their perception of the elements of the road (5). Furthermore, with the advent of virtual reality (VR) technology, application of bio-sensors is
no more limited to naturalistic studies. A recent bicycling behavior study explored participants’
engagement levels in immersive 360-degree using different movement controllers by means of an
electroencephalography (EEG) sensor (6). Another bicycling behavior study examined the ecological
validity of immersive and non-immersive presentation methods (7). The results demonstrated
the ability of VR to elicit behavioral patterns in line with those observed in real-world which was
also characterized by a higher degree of engagement.

EDA sensors have been used in transport studies to identify underlying emotional affects
of road users under various environmental circumstances. EDA measurements have been used to
investigate the influence of unpredictable events, traffic volume, and road design on car drivers
and passengers (8), to identify stressful events while driving (9), to inspect discomfort situations
in automated driving (10), and to evaluate drivers’ trust in automated vehicles (11). EDA sensors
have also been applied to understand pedestrians’ emotions and perceptions (12), to study objective
characteristics of the environment on walkability (13), and to measure social repulsive force based
on the human physiological response to stress (14).

The study at hand employs the immersive 360-degree virtual reality (VR) method combined
with an EDA sensor to study bicycling stress level in different infrastructure layouts. VR supports
a first-person experience of riding in currently available and non-existing infrastructure, where
participants can be safely surveyed. Besides, the controlled laboratory setting of VR experiments
allows for the application of sensitive physiological sensors which provide a deeper understanding
of hazardous encounters and stressful situations at a finer level of granularity. Namely, the current
study investigates the hypothesis that a significant difference exists between the average arousal
level of the bicyclists when riding in different environments.

The remaining part of the paper is organized as follows: Section “Literature Review” sum-
marizes previous studies using EDA sensors in bicycling research. The section “Methodology and
materials” introduces the proposed methodology to study bicyclists’ stress level and the materials.
Section “Analysis method” explains the analysis method of EDA data in detail, focusing on the
key performance measures. The study results are presented in section “Results” followed by the
discussion of the results in section “Discussion & Outlook”.

LITERATURE REVIEW

Various psychological processes such as cognitive load or stress activate the sympathetic branch of
the nervous system which results in the activity of the skin sweat glands (9, 15, 16). The increase
in the skin conductivity is due to sweating as an electrolyte solution that causes changes in the skin
resistance. The skin conductivity can be measured by passing a small electrical current across two
electrodes placed on the surface of the skin that measures momentary changes in the level of the
sweat gland activity. The amount of current that passes between the electrodes is defined as skin
conductance (measured in Micro Siemens).

Different types of physiological sensors, including electrocardiograms, electromyograms,
skin conductance, and respiration, have been used to determine drivers’ stress level during real-
world driving tasks (9). Skin conductance found to be the most accurate real time correlate of
the drivers’ stress, followed by HR and heart rate variability (HRV). Even though paralleled peaks
were observed between EDA and HR bio-signals reflecting mental workload and task determined
arousal patterns (17), the superiority of electrodermal variables over cardiovascular variables with
respect to persistent emotional strain during aviation experiments has been proven (18).

Skin conductivity is one of the most robust non-invasive physiological measures of auto-
nomic nervous system (ANS) activity which reveals different aspects of closely related emotional states and stress (19). Advances in digital recording of physiological sensors have allowed their application in ambulatory environments, laboratory-based experiments (15), as well as real-life studies (9) in different disciplines, in particular engineering psychophysiology.

In addition to studying drivers and pedestrians, EDA sensors have also used to study the arousal level of bicyclists (20–23). Zeile et al. (20) used a wristband to measure skin conductance, electrocardiograms, skin temperature, and HRV of bicyclists to identify moments of stress. Their findings showed that this approach could pinpoint situations with emotional peaks, particularly fear and anger. Berger and Dörrzapf (21) identified stressful events on a bicycle ride by means of an EDA sensor and an instrumented bicycle. They found an increase in EDA when bicyclists rode against the wind, yet declared that it could be due to an increase in sweating, caused by putting more effort to ride against the wind. However, a major drawback of such field experiments is lack of control over the environmental factors.

Jones et al. (22) attempted to identify and measure points of emotional arousal (e.g. stress and relaxation) along the routes, by means of wearable EDA sensors and equipped bicycles with action cameras and GPS devices. Many instances such as unpredictable pedestrian activity raised levels of stress and situations in which participants felt uncertain about what they were expected to do, such as negotiating junctions, produced feelings of anxiety, frustration and vulnerability. However, the EDA data was not normalized in this study, which is a prerequisite of data aggregation since EDA readings of each participant have a different baseline and response range and are, therefore, not directly comparable.

Caviedes and Figliozzi (23) applied an EDA sensor to observe the stressful events when bicycling in different traffic conditions, bicycle facilities, and intersections. By matching videos with stressful events, they observed the circumstances of stressful events and assumed that any changes in skin conductance were related to stressful events. EDA data was normalized according to Healey (24) to address the between-subject variation of baseline EDA along with random effects model to take into account individual effects and serial correlation. Based on their real world application, bicyclists experienced more stress levels during peak hours and also at signalized intersections, while separated bicycle infrastructure showed the lowest stress levels. However, a major drawback of this study was the placement of the sensor electrodes on fingers, which were in close contact to handlebar and brake levers. Hand movements (especially when pulling the brake levers) could compress electrodes and could interfere with and distort signals while riding a bicycle.

Nuñez et al. (25) presented a method for monitoring stress to analyze the influence of noise, vibration, presence of bicycle paths, and the period of the day on stress experienced by bicyclists. They generated maps which identified critical points of stress along routes. A logistic regression model revealed that high levels of noise increased the odds of experiencing stress by 4%, intensive vibrations increased the odds by 14%, and the presence of bicycle paths reduced the odds by 8%. It was also found that the odds of having stress increased by 24% in the afternoon rush hour compared to the morning rush hour.

Kyriakou et al. (12) examined the application of EDA sensors for outdoor experiments by conducting laboratory-based experiment as well as real-world field study to ensure transferability between laboratory settings and real-world field studies. They proposed an algorithm based on EDA, skin temperature, self-reported subjective perceived safety at different locations, and recorded ego-perspective videos by the participants to detect moments of stress. Their results
showed that the proposed algorithm detects moments of stress with 84% accuracy and high correlations between the data sources. In their pilot study they investigated whether a wider multi-purpose bike lane would improve bicyclists’ safety by probing participants’ emotional states. All of the data sources confirmed that the most common stress triggers are cars passing by closely, long waiting times, parked cars on bike lanes, and construction areas.

Werner et al. (26) investigated stress sensations of bicyclists through quantifying physiological measurements. The identified moments of stress were aggregated for all participants to identify more stressful elements of the road. The authors recommended a more detailed analysis at the level of subsections of a route instead of a global assessment of a trip, incorporating traffic volume and other factors which might influence stress levels, and surveying a larger sample for further research. They also suggested decomposition of skin conductance signal under non-laboratory conditions, a topic which will be reflected more in detail in Section 3.

METHODOLOGY AND MATERIALS

Materials

In this study, use was made of a bicycling simulator combined with immersive virtual reality. Participants were able to brake and pedal. Five different bicycling environments were created in VR: sidewalk next to pedestrians, painted bicycle path on the sidewalk, painted bicycle path on the road, roadside next to vehicles, and segregated bicycle path. Short videos are available online 1. Details of the virtual environments and the bicycle simulator can be found in Nazemi et al. (27) and Schramka et al. (28), respectively.

The Shimmer3 GSR+ sensor, which has been widely used in different research domains 2, was used to monitor the emotional status of each participant when riding in VR 3. This device is capable of measuring the electrical characteristics or conductance of skin, as well as capturing an optical pulse / Photoplethysmogram (PPG) signal and converting it to estimate heart rate (HR), using an ear clip or optical pulse probe. This sensor is non-invasive consisting of a small plastic box containing the hardware and electrodes that are attached to the skin.

Experimental design

A 5 x 2 mixed design was used with bicycling environment as within-subject factor and pedestrian / traffic volume as between-subject factor. Two different sequences for the bicycling environments were designed to account for learning and ordering effects.

– Sequence 1: sidewalk, painted bicycle path on the sidewalk, painted bicycle path on the road, roadside, segregated bicycle path, and
– Sequence 2: segregated bicycle path, roadside, painted bicycle path on the road, painted bicycle path on the sidewalk, sidewalk.

The Shimmer3 GSR+ unit was fastened to the arm with two electrodes connected to the inner wrist of the non-dominant hand of the participants to read the EDA. Normally, the palm side of the hand is used, however, it was found that grabbing the handlebar increased pressure on the electrodes and distorted the signal. It has been also recommended to use the non-dominant hand which generally has less movements to avoid artifacts (19). Accordingly, only right-handed partic-

1 Videos of the five different bicycling environments: https://youtu.be/ta8mAoDvyPY
2 http://www.shimmersensing.com/support/publications/
3 See https://youtu.be/PgJ9jFgJRBQ
Participants were invited to the experiment. In addition, the brake lever on the left side of the handlebar was taken out to reduce any unnecessary movements of the hand due to braking. Cleaning the skin with alcohol swabs before attaching the electrodes was considered, but it was found that it disturbed the signal quality. Therefore, it was decided to continue with the normal skin oil. EDA data was sampled at 16 Hz. The building air conditioning maintained the room temperature at around 24 °C (75.2 °F) for all participants as well as controlled the humidity of the room.

Experiment protocol

The findings in this paper derive from a larger experimental research project investigating bicyclists’ perceived safety. A researcher facilitated the entire process of the experiment for each participant. After reading an information sheet and signing a consent form, the EDA sensor was attached to the wrist of each participant. The study consisted of six parts; in the first part of the study participants were asked to read a short story to determine their EDA activity in normal conditions and to determine a baseline. The start and end time of reading the story was tagged by the researcher which was later used for analysis. Upon finishing the story, participants continued to the second part and answered questions about their socio-demographics, bicycling behavior, and attitudes towards bicycling on a desktop computer. Following the second part, participants were immersed in VR to check the details of the 360-environment for one minute, which was named “Orientation 1”. Later in the third and fourth parts of the experiment, participants were asked about perception of speed and space in VR. These parts were followed by “Orientation 2” or test course where participants were to learn bicycling in VR by reaching the finish line in an empty scene, i.e. no pedestrians or vehicles. Only in the fifth part participants bicycled in five different environments in VR with other road users present. Participants took off the head-mounted display and got off the bicycle to answer to a number of qualifying questions on the computer after each scene. In the last part of the experiment, participants were asked about their general experience with VR. The experiment protocol is shown in Figure 1.

“Orientation 1” was designed to reduce participants’ distraction due to curiosity about the details of the 360-degree scene. Therefore, participants were asked to check every corner of the scene, their bicycle, and their avatar in VR. In addition, this phase taught them to move their head freely in VR, which later helped them to confidently turn their heads at the intersection and check for the turning vehicles. The intention of “Orientation 2” was twofold: first, to teach participants how to bicycle in VR and familiarize them with the bicycling course, and second, to neutralize any arousal due to the anxiety of bicycling in VR for the first time that could affect their EDA data, which was recognized during the pilot tests.

Participants

Participants were mainly recruited from the National University of Singapore and the Land Transport Authority (LTA) of Singapore. Student participants were informed by an advertisement posted in the university website. LTA employees were invited to participate by sending an e-mail explaining the experiment held at the LTA campus. Students were compensated with S$15 cash and LTA employees were compensated with S$15 vouchers for about 45 minutes of their time. Participation eligibility criteria were as follows:

- Singaporean or permanent residents,
- Age between 18 and 65,
- Right-handed,
FIGURE 1 Experiment protocol

1. No acute physical or mental disorders, and
2. Some bicycling experience.

3. **Skin conductance**

   EDA signal is composed of an overall slow drifting signal, called tonic level, overlaid by short-term phasic fluctuations, called skin conductance responses (SCRs). The tonic level describes the overall level of EDA across time windows of ten or more seconds and is a suitable measure for comparing experimental tasks with long stages. The phasic activity describes a more spontaneous response which occurs when encountering specific abrupt, short-lived stimuli Figner and Murphy (29). Such response is the result of emotional arousals, i.e. a negative or positive affective valence. A response may happen due to external factors such as exposure to a sudden sound or a change in lighting, as well as internal factors such as formulating mental plans or having thoughts of expectation Damasio (30). There are various methods to analyze SC data (refer to Boucsein (19), Society for Psychophysiological Research Ad Hoc Committee on Electrodermal Measures (31), and Posada-Quintero and Chon (32) for more information). The deconvolution approach proposed by Benedek and Kaernbach (33) and continuous decomposition algorithm was used for SC data processing in this research.

   Extracting parameters of the EDA signal is required prior to conducting any statistical analyses. Each SCR is characterized by a number of features that are obtained after processing skin conductance data. These features include latency, amplitude, rise time, and half-recovery.
time. A subset of these measurements are usually used in each study for the assessment of the results. Figure 2 shows an ideal SCR to a hypothetical stimulus. The response takes place a few seconds after the stimulus. The interval between stimulus onset to response initiation is called latency and is normally between 1 to 3 seconds Dawson et al. (34). Amplitude is the phasic increase in SC from SCR onset to its peak and is the most commonly used measure. Rise time is the temporal interval between SCR initiation and SCR peak and half-recovery time is the temporal interval between SCR peak and point of 50% recovery of SCR amplitude.

![Graphical representation of principal SCR features](image)

**FIGURE 2** Graphical representation of principal SCR features  
Dawson et al. (34), Figure 10.5.

8 ANALYSIS METHOD
9 EDA processing
10 Pre-processing
11 It is required to visually inspect the EDA signals to identify artifacts, even if and automatic function has been applied on the data, because certain factors that were not predicted might be present in the data Braithwaite et al. (35). Therefore, each participant’s signal was visually checked to correct any artifacts. The artifacts arose mainly due to hand movements when wearing and taking off the head-mounted display or movement of electrodes on the wrist. There were very few cases where artifacts had happened during bicycling, as the participants were asked to keep their hands on the handlebar steady. The next stage of was the extraction of the phasic and tonic components from the recorded signal. Ledalab package was used for signal decomposition in MATLAB. Details of the continuous decomposition algorithm used in this package can be found in Benedek and Kaernbach (33). Initially, a threshold of 0.01 was selected for SCR detection to identify all SCRs above this value.

Smoothing methods were applied on the entire dataset to remove small shakes in the signal, which largely corresponded to low-pass filtering of the data. This pre-processing step is crucial in addition to artifact correction, especially when the number of detected skin conductance responses (nSCR) is to be used as a performance measure. Due to the long and non-stationary behavior of the signals, a Gaussian filter was applied with a window size equal to 48, which is three times the data collection frequency of 16. Visual inspection of the results revealed that a window size of 48 performed better in identifying SCRs while reducing significant number of irrelevant peaks due to small shakes in the signal, particularly in half recovery time of SCRs, which otherwise would result...
in falsely detected SCRs. In addition to the visual inspection, a sensitivity analysis was performed for the window size of the Gauss method. The smoothing was performed by increasing the window size from 16 to 32, 48, and 64 to observe the changes incrementally. Similar conclusions were derived by checking the marginal changes in the nSCR and mean SCR, which showed how a window size of 48 outperformed the other window sizes.

The SCR threshold was historically reported at the value of 0.05 µS, which was approximately the smallest shift visible on paper chart recorders. However, advances in technology and accuracy of sensors have led to capturing small skin conductance variations to the level of 0.01 µS, which is found more often in the recent literature (35). Although the majority of detected SCRs fell between 0.01 µS and 0.05 µS, due to the dynamic nature of the experiment that involved physical activity and body movements, the 0.03 µS was selected as the threshold for statistical analysis, which is also the recommended threshold (35).

The expected order of amplitude of SCRs depends considerably on individuals and experimental situations. Nevertheless, typical SCR amplitudes is roughly estimated to reach around 2 µS to 3 µS from threshold. For experiments that include highly aversive or fearful stimulus, the maximum response can rarely increase to 8 µS (35). Even though the limited extreme SCR values for this experiment could be considered justifiable and within the acceptable range (if they were not recognized as artifacts), they eventually were treated as outliers and SCR data was capped at the 99.5 percentile (= 3.02 µS) to bring the SCRs within the expected range. The truncation process affected 92 records in total with the maximum detected SCR of 7.94 µS that were reduced to 3.02 µS. In total, 9351 SCRs were identified for all 89 participants with a mean nSCR of 105 per individual and standard deviation of 49. Truncating the data reduced variance in the calculations of the means which were later used to perform statistical comparisons.

### Inter-individual variance reduction

Several transformation methods have been suggested for the SCR amplitude, due to inter-individual variations in nSCR and response range, prior to performing any statistical analysis (36). A transformation into standard values was selected for this study, which considers each particular individual mean and standard deviation of SCRs. Standardized SCRs are calculated as in Eq. (1).

$$
\begin{align}
  z_{ik} &= \frac{SCR_{ik} - \bar{SCR}_i}{s_i} \\
  T_{ik} &= 50 + 10z_{ik}
\end{align}
$$

where

- $z_{ik} = k$th SCR standardized value for individual $i$,
- $SCR_{ik} = k$th raw SCR score for individual $i$,
- $\bar{SCR}_i = \text{mean of all SCRs for individual } i$, and
- $s_i = \text{standard deviation of all SCRs}$.

The normally distributed z scores are transformed to T scores with a mean of 50 and standard deviation of 10 to drop out the minus signs as in Eq. (2) (19).
where
\[ T_{ik} = k \text{th SCR T score for individual i, and} \]
\[ z_{ik} = k \text{th SCR standardized value for individual i}. \]

**RESULTS**

**Sample: descriptive statistics**

The experiment was completed by 150 participants. Due to a drop in the Bluetooth connection of the bicycle simulator sensors, the data for 6 more participants got lost. Initially, the Empatica E4 EDA sensor was used for the first 40 participants. Unfortunately, the device was not reliable and there was a considerable failure rate of 38%. Even though this sensor was properly attached to the skin, in general, signal amplitude was low and for several cases the device had failed to collect any data. This is in line with recent research that also found discrepancies between the results of this device compared to a clinical laboratory-grade device, that could root in the low Empatica E4 sampling rate of 4 Hz Borrego et al. (37).

Sensor failure in the present research led to exclusion of 40 observations from the analysis. The Shimmer3 GSR+ was used for the rest of the participants, which robustly collected data with a reasonable failure rate of 14% mainly due to the dry skin type of the participants or a loose connection of the sensor electrodes to the skin. It has been found that approximately 10% of participants in any experiments are estimated to be non-responders (hypo-responsive) in terms of their EDA Braithwaite et al. (35).

Overall, there were 89 participants with proper data that were used for analysis (\( M_{\text{age}} = 28.8, \text{SD}_{\text{age}} = 9.5 \)). The majority of the participants’ were aged between 18 to 34. Bicycle was being used as a daily commute mode only by 3% of the participants and 63% never commuted by bicycle. All participants rode through all five environments with 42 and 47 participants experiencing high and low pedestrian / traffic volumes, respectively.

**Environment-level analysis**

Table 1 summarizes the mean T scores of the SCR amplitudes for different bicycling environments. Repeated measures ANOVAs were conducted on EDA data with bicycling environment as within-subject factor and one-way ANOVAs were applied on EDA for ambient traffic level as between subject factor. Filtering the SCR amplitudes caused unavailability of responses for some bicycling environments for a few participants. Therefore, a linear mixed effects model was replaced for the analysis to avoid dropping participants with missing T scores. Results showed that there was no significant difference between any pairwise comparison of the bicycling environments. One-way ANOVA also did not find any significant differences between mean T scores of low volume scenarios versus high volume scenarios.

**Segment-level analysis**

Bicycling environments were divided into six discrete segments as shown in Figure 3, such that the arousals within each environment can be studied in more detail. The segmentation was based on the average bicycling speed of participants; start segment was mainly used for acceleration to reach the desired bicycling speed. Participants maintained their desired speed in segment 1 and segment 2 with quite uniform speeds for each environment. Before intersection segment was where the
## TABLE 1 Mean and standard deviation values of T scores in each environment

<table>
<thead>
<tr>
<th>Sequence and ambient pedestrian / traffic volume</th>
<th>Number of participants</th>
<th>Descriptive statistics</th>
<th>Sidewalk Painted bicycle path on the sidewalk</th>
<th>Painted bicycle path on the road</th>
<th>Side</th>
<th>Segregated bicycle path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence 1 low pedestrian / traffic</td>
<td>22</td>
<td>M (SD)</td>
<td>52.81 (11.72)</td>
<td>49.47 (9.53)</td>
<td>49.32</td>
<td>51.13 (10.78)</td>
</tr>
<tr>
<td>Sequence 1 high pedestrian / traffic</td>
<td>16</td>
<td>M (SD)</td>
<td>50.53 (10.15)</td>
<td>49.86 (9.76)</td>
<td>50.70</td>
<td>52.23 (12.48)</td>
</tr>
<tr>
<td>Sequence 2 low pedestrian / traffic</td>
<td>25</td>
<td>M (SD)</td>
<td>49.82 (9.94)</td>
<td>50.83 (10.22)</td>
<td>50.07</td>
<td>50.87 (10.37)</td>
</tr>
<tr>
<td>Sequence 2 high pedestrian / traffic</td>
<td>26</td>
<td>M (SD)</td>
<td>50.95 (10.14)</td>
<td>50.91 (10.51)</td>
<td>50.88</td>
<td>49.87 (9.77)</td>
</tr>
<tr>
<td>Aggregated data</td>
<td>89</td>
<td>M (SD)</td>
<td>51.25 (10.73)</td>
<td>50.19 (9.98)</td>
<td>50.18</td>
<td>50.97 (10.82)</td>
</tr>
</tbody>
</table>

Intersection became visible to the participants in VR and they started to decelerate. Participants accelerated at the intersection segment to cross the intersection until they reach their desired speed. Obviously, finer level of segmentation can be considered for more detailed analysis.

The mean T score of each segment was calculated by averaging the SCR T scores of all participants that occurred in that segment. The mean SCR T scores of each segment is presented in Figure 3 with their relative standard deviations expressed in parenthesis. It was observed that participants often had an initial arousal when immersed in new bicycling environments, which happened in the start segment. Furthermore, participants were anticipating the finish line during the end segment and got excited by passing the finish line which caused some arousals. Therefore, this segment was excluded from the analysis of bicycling environment. Sections Before intersection and intersection clearly have higher mean SCR T scores within each environment which reflects the amount of stress participants incurred to cross the intersection. It is interesting how this stress existed even in the empty test course scene as the first scene, although participants had been primed there will be no other people. The mean SCR T score of segment before intersection for the roadside environment had the highest value followed by segregated bicycle path; this high value for the segregated bicycle path could be due to the fact that participants suddenly found themselves unprotected when arriving at the intersection, which caused sudden responses.
## FIGURE 3

Mean and standard deviation of SCR T scores at each segment of bicycling environments.
It is believed that segment 1 best represented the influence of bicycling environments on the stress level of the participants; during this segment, participants had already gained enough familiarity with the environment after finishing the start segment, and tried to maintain their desired speed in segment 1. It can be seen how bicycling on the sidewalk was more stressful due to unexpected movements of pedestrians that could potentially come to the bicycling path of the participants. This concern still existed, but to a lesser extent, on the painted bicycle path on the sidewalk as well as the roadside.

Habituation phenomenon—a from of learning process characterized by decreasing response intensity to a stimulus after repeated or prolonged presentations of that stimulus—showed its effect in segment 2, where participants had already learned about the behavior of other road users in the previous segments. Therefore, the arousals in this segment were less pronounced compared to segment 1, due to any common stimulus. Furthermore, since the occurrence and frequency of the events were random, the mean SCR T scores found to be different from segment 1. The highest value of mean SCR T score for segment 2 belonged to roadside bicycling, in which many motorized cars passed by the participants along this long stretch of road.

The mean SCR T scores for the before intersection segment found to be lower for painted bicycle path on the sidewalk and sidewalk. The design of the road allowed the participants to easily have the turning cars in their view while bicycling on the sidewalk. Furthermore, the fact that bicyclists were riding next to the pedestrians might have made them feel less stressed expecting the drivers to yield for them. On the contrary, when bicyclists were riding closer to the motorized vehicles, even though there were no barriers between participants as bicyclists and motorized traffic providing good visibility to the drivers, participants were more stressed as they needed to fully turn their heads to check for turning cars.

Further statistical analysis was conducted to explore any differences in sensor measures due to the environment characteristics. To this end, repeated measures ANOVA with random effects was conducted to compare segment 1 of all environments. Results showed that SCR T scores of segment 1 were significantly different from each other ($p < 0.05$). Tukey HSD post-hoc analysis revealed that segment 1 of the segregated bicycle path had significantly lower SCR amplitudes compared to segment 1 of the sidewalk ($p < 0.05$). Similar analysis was conducted for segment 2. More significant differences were found in this segment. Segment 2 of all environments had significantly lower SCR amplitudes compared to segment 2 of sidewalk ($p < 0.001$). Highest differences considering the order were related to the segregated bicycle path, the painted bicycle path on the road, the roadside, and the bicycle path on the sidewalk. In terms of the objective measurements of bicycling stress, roadside bicycling performed better than the sidewalk. One implication could be the number of conflicts with pedestrians and imposed stop and goes to the bicyclist that occurred frequently on the sidewalk.

A mixed effects model was used to have a between-segment comparison of the T scores. Table 2 shows the regression results with Segment 1 as the reference segment; segment 1 was expected to elicit the most natural responses of participants from any of the environments and had the least estimated T score as well. T scores were significantly higher in Before intersection followed by Intersection compared to Segment 1 ($p < 0.001$), which reflects how participants got stressed prior to crossing the intersection.
TABLE 2 Regression results for T scores of detected SCRs in different segments

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>SCR T score</th>
<th>Estimate</th>
<th>t-value</th>
<th>Sign.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td>48.98</td>
<td>171.81</td>
<td>***</td>
</tr>
<tr>
<td>Segment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Start</td>
<td></td>
<td>0.98</td>
<td>2.56</td>
<td>*</td>
</tr>
<tr>
<td>Segment 1 (reference)</td>
<td></td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Before intersection</td>
<td></td>
<td>4.78</td>
<td>11.50</td>
<td>***</td>
</tr>
<tr>
<td>Intersection</td>
<td></td>
<td>2.52</td>
<td>5.23</td>
<td>***</td>
</tr>
<tr>
<td>Segment 2</td>
<td></td>
<td>1.33</td>
<td>3.84</td>
<td>***</td>
</tr>
<tr>
<td>End</td>
<td></td>
<td>1.78</td>
<td>3.28</td>
<td>**</td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td></td>
<td>0.008</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sign. codes: 0 ‘****’ 0.001 ‘***’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

1 DISCUSSION & OUTLOOK

The present research discussed the results of the physiological sensor to measure bicycling stress in different environments. It sought to determine whether reliable estimates of bicycling stress could be obtained under virtual or simulated conditions; that is, by placing an individual within an immersive 360-degree virtual environment where bicycling environment, traffic volume, traffic speed, and curb lane width were all controlled. This study also explored the potential for applying this technology to engineering and behavioral aspects associated with bicycling and the design of new infrastructure.

It was assumed that bicyclists were likely to experience varying levels of stress and anxiety due to their individual perceptions of risk when bicycling in an environment with vehicular and pedestrian traffic. EDA analysis at the environment-level did not lead to significant differences between the environments. A study investigating the emotions of drivers via heart rate also showed that there was no significant difference between the heart rates when driving in three different road types of city, ring road, and motorway (38). However, segment-level analysis in the same study revealed that intersections were major points of stress as well as potential conflicts and collisions. Dozza and Werneke (39) similarly showed higher risk at intersections, which found to be in line with the findings of this research. Objective measurements of safety also showed that the majority of bicycle and motor vehicle collisions happen at intersections, compared to overtaking and dooring accidents (40).

A conflict is normally defined as an interaction between a bicyclist and other road users such that at least one of the parties has to change speed or direction to avoid a collision. Studies investigating these conflicts may offer valuable insights into how bicyclists and other road users behave during their interactions on various types of transportation infrastructure. However, it is not possible to determine whether the safety of the bicyclists was compromised during conflict events (41). The findings of this research demonstrated the changes in the arousal level of the bicyclists were affected by potential conflicts with pedestrians or motorized vehicles. Decomposing the EDA signal, as suggested by Werner et al. (26), and obtaining SCR values for different segments
allowed for the comparison of smaller segments of the environments. These individual segments were assumed to have different types of stimuli. It should be noted that the reported SCR rates for each segment were averages and the result of encounters with vehicular traffic and pedestrian traffic within each bicycling environment.

It was found that sidewalk bicycling was the most stressful of all environments, judging by the mean SCR T scores of the first segment. Furthermore, a relationship between the objective road assessments and participants’ stress level was confirmed based on the stress level of bicyclists on the sidewalk shared with pedestrians. Objective safety numbers has indicated that bicycling on the sidewalk is accompanied by higher accident rates; it has been shown for a study in Canada that the sidewalk bicyclists have higher event rates on roads than non-sidewalk bicyclists (42). The rates of injuries and crashes (collisions and falls) in bicycle-friendly cities of Toronto and Ottawa suggested that sidewalks are less safe compared to on-road bicycling (43). In line with this finding, similarly, studying safety critical events of 31 bicyclists riding for four weeks in a naturalistic study revealed that the majority of incidents are bicycle-bicycle and bicycle-pedestrian, rather bicycle-motorized vehicles, and they are mainly due to sudden unexpected maneuvers of other bicyclists and pedestrians (44). In addition, it has been discovered that considering the potential conflicts between pedestrians and sidewalk bicyclists, the number of collisions between these two road users has been very limited and there are larger number of sidewalk collisions between the bicyclists themselves (42). As such, the low accident rate of sidewalk bicycling would not correctly indicate whether sidewalk bicycle facilities perform safely or not. It can be concluded that the policies which encourage or tolerate sidewalk bicycling should be amended and the solution to sidewalk safety is to remove the bicyclists from the sidewalk.

EDA sensor data also indicated how a segregated bicycle path outperforms in reducing bicycling stress levels compared to other types of bicycle facilities. Therefore, an important step to be taken in improving the bicycling stress of bicyclists would be the introduction of more dedicated bicycle facilities which protect bicyclists from motorized vehicles. Depending on the road type (i.e residential, commercial, etc.), traffic volume and prevailing speed limits, painted bicycle path on the road or even integration of the bicyclists and motorized traffic can be promoted through the implementation of wide outside lanes. Moreover, design of any bicycle facility on roads with many intersections might not be attractive to bicyclists, due to the imposed stress at every intersection. A suitable design might be similarly integration such that the bicyclists consider themselves as part of the traffic stream and at the same time visible to the motorized vehicles. If sidewalk bicycling is encouraged for any reason in the aforementioned road, clear sight lines at intersections must be considered by urban roadway and bicycle infrastructure designers to avoid blind conflicts.

AUTHOR CONTRIBUTION STATEMENT
Mohsen Nazemi contributed to the experimental design, executed the traffic microsimulation, conducted data collection, performed the descriptive and statistical analysis, interpretation of the results, and has written the manuscript. Dr Michael van Eggermond contributed to the experimental design, provided input on the descriptive and statistical analysis, interpretation of the results, and has written the manuscript. All authors reviewed the results and approved the final version of the manuscript.
ACKNOWLEDGEMENTS
This research has been carried out at the CREATE campus of the Singapore ETH Centre and was funded by the Singaporean National Research Foundation (NRF). The authors appreciate the consultation of Dr. Panos Mavros, Valentin Melnikov, and Dr. Victor Schinazi regarding the analysis of the physiological sensor data, which led to insightful ideas. The authors would like to acknowledge Michael Joos for the game development and Unity integration, Tanvi Maheshwari for street design and Dorothea Schaffner for her support with the survey and experimental design. The authors would also like to thank the Land Transport Authority (LTA) for their support and the possibility to conduct the study on their premises.

REFERENCES


