Data in Brief 57 (2024) 110968

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Mobile brain–body imaging data set of indoor treadmill walking and outdoor walking with a visual search task

Grant M. Hanada[a], Marija Kalabic[a], Daniel P. Ferris[b,*]

[a] University of Michigan, Ann Arbor, MI, USA
[b] University of Florida, Gainesville, FL, USA

ARTICLE INFO

Article history:
Received 8 July 2024
Revised 17 September 2024
Accepted 19 September 2024
Available online 21 September 2024

Dataset link: LSIE individual subjects full set (Original data)
Dataset link: MOBILE BRAIN-BODY IMAGING DURING INDOOR TREADMILL WALKING AND OUTDOOR OVERGROUND WALKING WITH A VISUAL SEARCH TASK (Original data)

Keywords:
EEG
Brain
Real world
ERP
Gait
Kinematics
Heart rate
Virtual reality

ABSTRACT

To fully understand brain processes in the real world, it is necessary to record and quantitatively analyse brain processes during real world human experiences. Mobile electroencephalography (EEG) and physiological data sensors provide new opportunities for studying humans outside of the laboratory. The purpose of this study was to document data from high-density EEG and mobile physiological sensors while humans performed a visual search task both on a treadmill in a laboratory setting and overground in a natural outdoor setting. The data set includes 49 young, healthy participants on an outdoor arboretum path and on a treadmill in a laboratory with a large virtual reality screen. The data provide a valuable research tool for scientists interested in signal processing, electrocortical brain processes, mobile brain imaging, and brain-computer interfaces based on mobile EEG. Given the comparison data between laboratory and real world conditions, researchers can test the viability of new processing algorithms across conditions or investigate

* Corresponding author.
E-mail address: dferris@bme.ufl.edu (D.P. Ferris).
Social media: @Ferris_vball (D.P. Ferris)

https://doi.org/10.1016/j.dib.2024.110968

changes in electrocortical activity related to behavioural dynamics coded into the data.

Specifications Table

Subject	*Neuroscience: General*
Specific subject area	*Human mobile brain/body imaging*
Type of data	*High-density electroencephalography, inertial measurement unit, treadmill ground reaction forces, heart rate, eye tracking, handheld button response, and video as time synchronized time series. Salivary cortisol as analysed discrete data. Raw and Processed*
Data collection	*256-channel Biosemi Activetwo EEG; 8-channel neck muscle Biosemi EMG, 6 APDM Opal inertial measurement units attached to both feet, both ankles, waist, and chest; Bertec split-belt treadmill; Pulse Sensor heart rate monitor, Mobile Eye XG eye tracker, handheld Canon camcorder, handheld electronic button, and SalivaBio Oral Swab salivary cortisol samples*
Data source location	*University of Michigan, Ann Arbor, MI, USA*
Data accessibility	*Data are stored in IEEE Dataport* (https://ieee-dataport.org/open-access/mobile-brain-body-imaging-during-indoor-treadmill-walking-and-outdoor-overground-walking; DOI: 10.21227/H24T0V) *as an open-access data set. A second copy is stored as 49 individual .zip files by subject is at Figshare* (https://doi.org/10.6084/m9.figshare.6741734; DOI: 10.6084/m9.figshare.6741734) Repository name: *1) IEEE Dataport, 2) Figshare* Data identification number: *1) DOI:* 10.21227/H24T0V, 2) DOI: 10.6084/m9.figshare.6741734 Direct URL to data: *1)* https://ieee-dataport.org/open-access/mobile-brain-body-imaging-during-indoor-treadmill-walking-and-outdoor-overground-walking, 2) 10.6084/m9.figshare.6741734 Supplementary data set of gait events matched with EEG data sets:. 10.6084/m9.figshare.27022414.v1
Related research article	*None*

1. Value of the Data

- The data provide a comparison of similar data collected in an outdoor real-world locomotion setting and in a constrained laboratory locomotion setting.
- Data can allow for testing of novel signal processing algorithms for removing noise and testing hypotheses about brain function.
- Physiological data simultaneously collected with electrical brain data provide insight into brain-body connections.

2. Background

Identifying neural correlates and their dynamics in ambulatory tasks have presented many challenges in research. The vast majority of neuroimaging studies on human brain activity have been confined to laboratory environments. The overall goal of this project was to document the possibility of high-density electroencephalography (EEG) to provide insight into human brain function during ambulatory tasks in real world environments. The current experimental data collection examined young, healthy participants walking in a natural arboretum environment while they wore a high-density (256 channel) EEG. Subjects completed a visual search task where they had to identify coloured flags in their environment. Subjects also completed an indoor virtual reality version of the task on a treadmill. By completing similar tasks indoors in a laboratory on a treadmill, and outdoors in an arboretum on natural terrain, the data set may reveal differences between traditional laboratory experiments and real world experiments. A secondary goal was to provide a potentially stressful intervention on the task, where subjects gained and lost monetary rewards based on their performance.

3. Data Description

The study data is set up as an EEG Study Schema (ESS) Standard Data Level 1 container. This means that it contains raw, unprocessed EEG data arranged in a standard manner. Data is in a container folder and ready to be used with MATLAB to automate access and processing. All other data measures other than EEG are in .mat (MATLAB) format. For more information please visit eegstudy.org.

There is one folder for every subject that includes the following files when available:

(1) Indoor EEG session (<ID number_Indoor.set>)

EEG files have been imported into EEGLAB and are stored as unprocessed raw .set format in standard EEGLAB Data Structures.

(https://sccn.ucsd.edu/wiki/A05:_Data_Structures)

(2) Outdoor EEG session (<ID number_Outdoor.set>)

Same as Indoor EEG session (above)

(3) Indoor IMU session (<ID number_Indoor_imu.mat>)

The IMU .mat file contains a structure with 6 fields (variable name: IMU)

IMU.dataLabel: string including ID number, environment, and sensor type

IMU.dataArray: 10xNx6 matrix. Third dimension refers to each of 6 IMU sensors (left foot, right foot, left ankle, right ankle, chest, and waist). Columns are frame numbers. Rows are:

- *x*, *y*, and *z* direction of accelerations, in m s^{-1}^2.
- *x*, *y*, and *z* direction of gyroscopes, in rad s^{-1}.
- *x*, *y*, and *z* direction of magnetometers, in microteslas.
- Temperature, in degrees Celsius.

IMU.axisLabel: String headings for ‘dataType’ and ‘frame’ and ‘sensorNumber’.

IMU.axisValue: 1×10 cell array of string headings for each row of data type, and 1×6 cell array of string headings for each IMU sensor

IMU.samplingRate: Sampling rate

IMU.dateTime: String of date and time information of recording

(4) Outdoor IMU session (<ID number_Outdoor_imu.mat>)

Same as Indoor IMU session (above).

(5) Indoor eye tracking session (<ID number_Indoor_eye_tracker.mat>)

The eye tracker .mat file contains a structure with 6 fields (variable name: Eye_tracker)

Eye_tracker.dataLabel: string including ID number, environment, and sensor type

Eye_tracker.dataArray: 7xN matrix. Columns are frame numbers. Rows are:

- *x* and *y* coordinates of the master spot, in eye image pixels.
- *x* and *y* coordinates of the pupil center, in eye image pixels.
- Pupil radius, in eye image pixels.
- Eye direction with respect to the scene image, in scene image pixels.

The eye and scene images are displayed and recorded with resolution of 640×480 pixels. The origin is the top left of the image with the X-axis positive to the right and the Y-axis positive downwards. Unavailable data is shown by the number –2000.

Eye_tracker.axisLabel: String headings for ‘dataType’ and ‘frame’

Eye_tracker.axisValue: 1×7 cell array of string headings for each row of data type

Eye_tracker.samplingRate: Sampling rate
Eye_tracker.dateTime: String of date and time information of recording

(6) Outdoor eye tracking session (<ID number_Outdoor_eye_tracker.mat>)

Same as Indoor eye tracking session (above).

(7) Indoor heart rate from pulse sensor session (<ID number_Indoor_pulse_sensor.mat>)

The pulse sensor .mat file contains a structure with 6 fields (variable name: Pulse_sensor)
Pulse_sensor.dataLabel: string including ID number, environment, and sensor type
Pulse_sensor.dataArray: 3xN matrix. Columns are frame numbers. Rows are:

- Pulse (normalized wave), in volts.
- Inter-beat Interval (IBI), in milliseconds.
- heart rate, in beats per minute (BPM).

Pulse_sensor.axisLabel: String headings for 'dataType' and 'frame'
Pulse_sensor.axisValue: 1 × 3 cell array of string headings for each row of data type
Pulse_sensor.samplingRate: Sampling rate
Pulse_sensor.dateTime: String of date and time information of recording

(8) Outdoor heart rate from pulse sensor session (<ID number_Outdoor_pulse_sensor.mat>)

Same as Indoor pulse sensor session (above).

(9) Indoor heart rate from EEG session (<ID number_Indoor_pulse_from_eeg.mat>)

If pulse rate was recovered from EEG ECG a corresponding file is available. The pulse from EEG .mat file contains a structure with 6 fields (variable name: Pulse_from_EEG)
Pulse_from_EEG.dataLabel: string including ID number, environment, and sensor type
Pulse_from_EEG.dataArray: 3xN matrix. Columns are frame numbers. Rows are:

- Pulse (normalized wave), in volts.
- Inter-beat Interval (IBI), in milliseconds.
- heart rate, in beats per minute (BPM).

Pulse_from_EEG.axisLabel: String headings for 'dataType' and 'frame'
Pulse_from_EEG.axisValue: 1 × 3 cell array of string headings for each row of data type
Pulse_from_EEG.samplingRate: Sampling rate
Pulse_from_EEG.dateTime: String of date and time information of recording

(10) Outdoor heart rate from EEG session (<ID number_Outdoor_pulse_from_eeg.mat>)

Same as Indoor pulse from EEG session (above).

(11) Indoor treadmill force plate session (<ID number_Indoor_force_plate.mat>)

The force plate .mat file contains a structure with 6 fields (variable name: Force_plate)
Force_plate.dataLabel: string including ID number, environment, and sensor type
Force_plate.dataArray: 3xNx2 matrix. Third dimension is for left and right force plates, respectively. Columns are frame numbers. Rows are:

- *x*, *y*, and *z* direction of force, in newtons.

Force_plate.axisLabel: String headings for 'dataType' and 'frame' and 'sensorNumber'
Force_plate.axisValue: 1 × 3 cell array of string headings for each row of data type
Force_plate.samplingRate: Sampling rate
Force_plate.dateTime: String of date and time information of recording

(12) EEG digitized head map (<ID number.sfp>)

Besa coordinates of all electrode positions.

(13) Indoor eye tracking video (<ID number_Indoor_eye_tracker.avi>)

The eye tracker .avi file is a video from the subject's perspective (640 × 480 resolution, 30 frames/s)

(14) Outoor eye tracking video (<ID number_Outdoor_eye_tracker.avi>)

The eye tracker .avi file is a video from the subject's perspective (640 × 480 resolution, 30 frames/s)

(15) Indoor video camera (<ID number_Indoor_video_camera(#).avi>)

The camcorder .avi file is a video from the experimenter's perspective (704 × 384 resolution, 30 frames/s). If there are multiple parts the (#) appended indicates the order.

(16) Outdoor video camera (<ID number_Outdoor_video_camera(#).avi>)

The camcorder .avi file is a video from the experimenter's perspective (704 × 384 resolution, 30 frames/s). If there are multiple parts the (#) appended indicates the order.

(17) Cortisol (Cortisol_all_subjects.xlsx)

Salivary cortisol data is provided as a single spreadsheet 'Cortisol_all_subjects.xlsx'. It contains the following variables:

subid: ID number
sex: 1 = male, 2 = female
age: in years
height: in inches
weight: in pounds
environment: 1 = outdoors, 2 = indoors
ordererenvironment: 1 = outdoor first, 2 = indoor first
orderstress: 1 = stress first, 2 = non-stress first
condition: 1 = Initial sample taken before walking started, 2 = Baseline sample after baseline walking, 3 = Non-stress sample taken after non-stress condition, 4 = Stress sample taken after stress condition
concentration: cortisol levels in µg/L

cond_ordered = order of conditions by environment

(18) Gait events (LSIE_<ID number>_<Indoor/Outdoor>_HED+Gait.mat)

Indoor (or outdoor) gait events and timings corresponding to each EEG indoor (or outdoor) session file. For indoor gait events, we used the treadmill forceplate z-axis force data to determine gait timings. For outdoor gait events, we used the inertial measurements units' accelerometer data that were placed on each foot to determine timings.

Each .mat file contains a structure with 4 fields (variable name: events)

events.type: String name for gait event (LHS = left heel strike, RHS = right heel strike, LTO = left toe off, RTO = right toe off)

events.latency: Integer sample number aligned with EEG.data structure. Sampling rate = 512 Hz. Example: latency = 512 corresponds to time = 1 s

events.hedtags: String name for any additional HED tag events

events.urevent: Integer index value for each ordered event number

4. Experimental Design, Materials and Methods

4.1. Subjects

Forty-nine healthy adults (20 males, 29 females) between the ages of 18–45 years (average age 22.7 years) participated in the study. None of the subjects had any history of neurological or physical impairments and were in good shape, such that they could walk on a treadmill and on outdoor terrain for one hour while carrying a fifteen-pound load without issue.

4.2. EEG acquisition

Subjects were fitted with a high density, pre-amplified 256-channel EEG cap (sampling rate: 512 Hz; Biosemi Activetwo, Amsterdam, Netherlands). We placed 8 additional sensors on the neck of the subject to record neck muscle activity. The position of each electrode relative to the subject's head was recorded using a 3D digitizer (Zebris, Germany). Before data collection, we applied electrode gel underneath each sensor and kept all offsets below 20 mV as recommended by Biosemi user manual for optimal data quality.

4.3. Inertial measurement units

Subjects wore 6 inertial measurement units (IMUs) (sampling rate: 128 Hz; APDM Opal, Portland, OR) attached to both feet, both ankles, waist, and chest. Each IMU monitored limb movements using 3D accelerometers, gyroscopes, and magnetometers.

4.4. Instrumented treadmill

Subjects walked on an instrumented, split-belt treadmill (sampling rate: 1000 Hz; Bertec, Columbus, OH) which has separate belts, each with its own motor, for the left and right side of the walking surface. Separate force transducers under each belt were used for the collection of 6 degrees of freedom ground reaction forces from the left and right foot separately.

4.5. Cortisol

The SalivaBio Oral Swab saliva collection method was used to obtain participant salivary cortisol samples (Salimetrics, Newmarket, UK). Subjects were previously informed not to consume any food or beverages (other than water) one hour prior to testing. Subjects were not allowed to eat during the duration of testing. A total of 8 saliva samples (4 indoors, 4 outdoors) were collected during testing as follows: before the start of each experiment, between the first and second conditions, between the second and third conditions, and after the third condition for both indoor and outdoor sessions. Collections were done by placing an oral swab under the subject's tongue for two minutes. Swabs were then removed and placed in storage tubes and held in an ice chest at below 4 °C during testing after which they were frozen at below −20 °C until analysis.

4.6. Heart rate

Heart rate was measured using photoplethysmogram (PPG) recordings obtained from a Pulse Sensor heart rate monitor (sampling rate: 500 Hz; World Famous Electronics LLC, New York, NY) attached to the subject's right earlobe and controlled through an external Arduino.

4.7. Eye tracker

Each subject had visual gaze recorded with Mobile Eye XG eye tracking glasses (sampling rate: 30 Hz; Applied Science Laboratory, Boston, MA) with 0.5–1° accuracy.

4.8. Video

High definition video recordings of the experiment were done using a high definition camcorder (sampling rate: 30 Hz; Canon USA) handheld by an experimenter.

4.9. Backpack

Subjects wore a backpack that contained the EEG amplifier, battery, eye tracker acquisition logger, and connectors such that a line of cables connected the multiple systems to the data collection laptop placed in a carrying tray held by an experimenter following behind the subject approximately 10 feet away (Fig. 1A and B).

4.10. Testing environments

4.10.1. Virtual reality environment

A 3D animated environment (designed in Google SketchUp) of a virtual park was displayed as a video on a projector screen directly in front of the treadmill (Fig. 2) set at 0.7 m s^{-1} moving at the same pace as the video. Within the virtual landscape, there were a total of 50 animated target flags (bright green) and 150 animated non-target flags (dark green) split evenly between the non-stress and stress conditions. There were no flags displayed during the baseline condition.

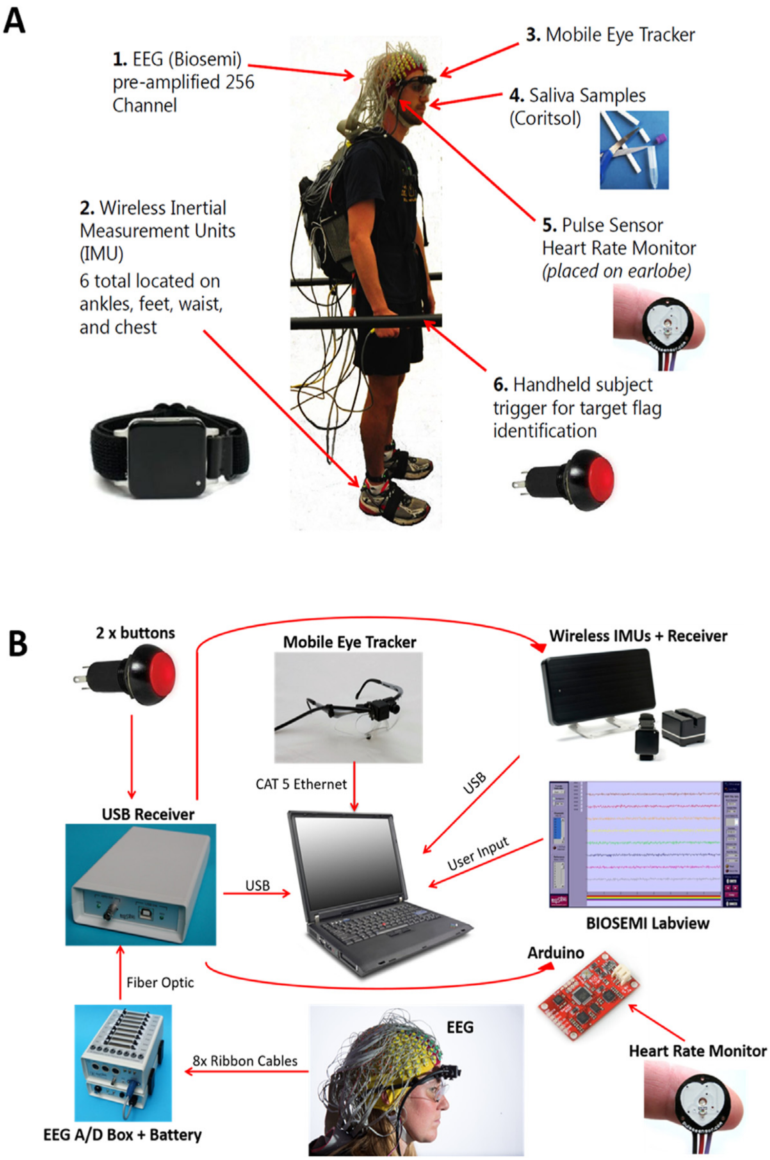

Fig. 1. (A) Diagram of measurement devices placed on an example subject. (B) Layout of recording devices and how they were connected to each other for synchronization.

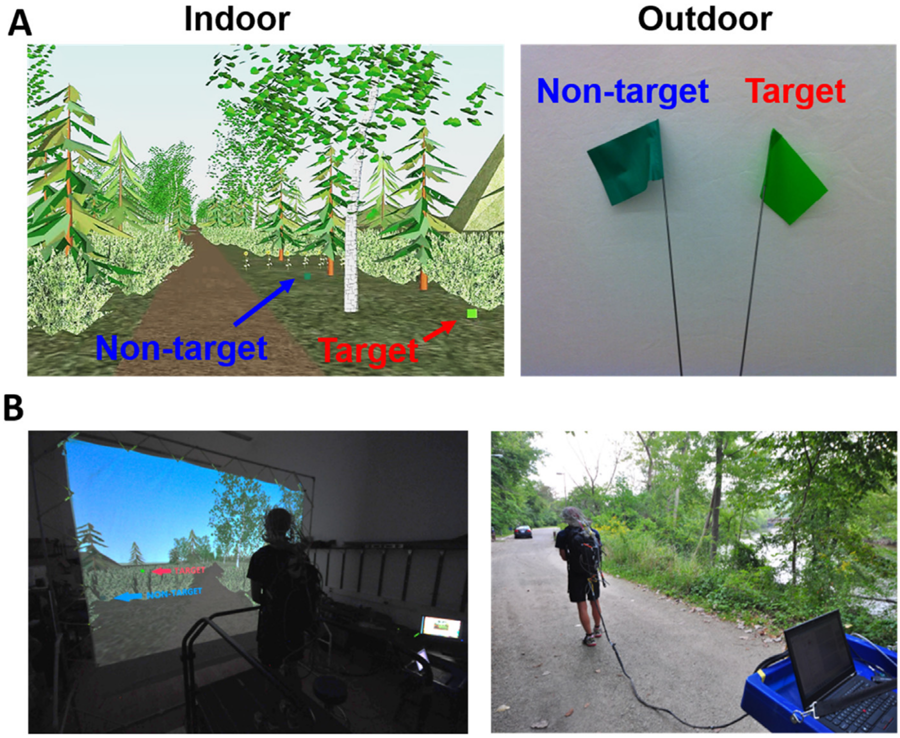

Fig. 2. Top two images in **A** show the stimuli used in the visual search tasks. The bottom two images in **B** show examples of data recording sessions. The bottom Left panel shows an example participant standing on the instrumented, split-belt treadmill in front of the projector screen displaying the indoor, virtual environment. The bottom right panel shows an example participant walking along the outdoor path tethered to an experimenter by cables from measurements devices connected to a laptop on a tray.

The difficulty of the target flag detection level varied based on position relative to environmental objects as well as distance from path.

4.10.2. Real world environment

Subjects walked outdoors on a marked trail path (Fig. 2) at an arboretum (Nichols Arboretum, Ann Arbor, MI). The path was accessible to the public during testing. The marked path was approximately 2 miles long in total (Fig. 3) and consisted mostly of hard packed dirt with some rocks/pebbles. There were a total of 50 target flags (bright green) and 150 non-target flags (dark green) split evenly between the non-stress and stress conditions. Flags were 2″ × 3″ trail marking flags attached to a 15″ pole. There were no flags in place during baseline walking. Subjects were instructed to walk straight along the path at a slow to moderate pace without stopping. Auditory cues to speed up or slow down were given if necessary. Flags were placed within 10 feet off the path and not higher than 10 feet off the ground. They were placed in such a manner that they were visible from the path without having to step off of it.

4.10.3. Protocol

Subjects walked for approximately one hour continuously in each environment while performing a visual search task. First, subjects started with a 20 min baseline condition of normal walking in which no flags were present in the environment. After baseline walking was complete the visual search task began. The goal of the task was to search and identify bright green flags (targets) within the environment and to ignore dark green flags (non-targets). Subjects were to

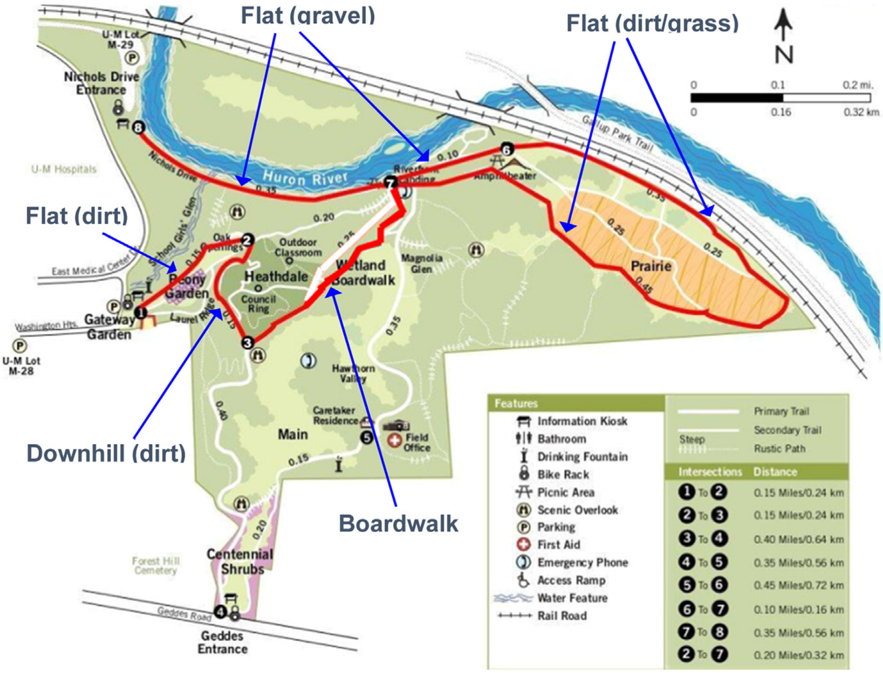

Fig. 3. Map of the arboretum with the walking path highlighted in red. Subjects started at maker 1, and at marker 7 continued northwest toward marker 6. After looping back to marker 7 subject proceeded west to the finish at marker 8. The predominant type of terrain is indicated on the map for each section between markers. However, the transitions in terrain were not always discrete, and reflected a continuum. The video files of each participant provides a clearer indication of how terrain was changing across the recording session as participants walked.

press the button on a joystick they were holding when they saw a target flag and to not press anything when they saw a non-target flag. There were 2 conditions, 20 min each, during the visual search task: 1) “Non-stress” – subjects were told each correct flag identified earned an extra $0.25 toward their compensation for the study and 2) “Induced Stress” - in addition to earning $0.25 per correctly identified target flag subjects were also told they would be penalized $1.00 for every unidentified target flag and a loud siren was played immediately after the subject walked passed the unidentified target flag. To further induce stress, subjects were given false negative siren noises indicating a missed target flag randomly (approximately once every 2 min, 10 times total). Subjects were naïve to the random penalties during the task, but were debriefed after the conclusion of the experiments. Baseline walking was always first, but the order of the 2 visual search conditions was randomized.

4.10.4. Code availability

Code for importing data measures from all sources into Matlab and EEGLAB is provided.

4.10.5. Data storage

All data measures used in this study are stored in IEEE Dataport (https://ieee-dataport.org/open-access/mobile-brain-body-imaging-during-indoor-treadmill-walking-and-outdoor-overground-walking; DOI: 10.21227/H24T0V) as an open-access data set (Data Citation 1). The data is stored as a single compressed .zip file (778 GB) and archived into individual folders by subject ID. A second copy archived as 49 individual .zip files by subject is also made available in Figshare (https://doi.org/10.6084/m9.figshare.6741734; DOI: 10.6084/m9.figshare.6741734) (Data Citation 2).

4.11. Technical validation

4.11.1. Data synchronization

All data collection devices (EEG, mobile eye tracker, IMUs, force plates, and heart rate monitor) were synchronized with custom software written in Labview to link with Biosemi (Fig. 1b). Synchronization triggers from each device were marked in Biosemi in order to align each device in time with EEG. After data collection, researchers confirmed synchronization on data files and that each device was synchronzied within 1–2 data frames of each given device's sampling rate. For camcorder video synchronization, during each EEG session we video recorded the experimenter's live computer screen containing the real-time data visualizations and timings such that in post we aligned the camcorder's clock time to the EEG timestamps and subtracted the offset between the two in order to be within 1 camera frame alignment accuracy (e.g. 1/30 s).

4.11.2. HED tagging

Hierarchical event descriptor (HED) tags are EEG semi-structured annotations used to provide categorized and detailed descriptors of events during an experiment (Bigdely-Shamlo et al., 2017). These events that are logged with a predefined, structured, and common annotated language across studies that use the HED system. Some of these events like button presses are tagged automatically in real-time during the experiment. But most other events were marked manually by an experimenter.

In order to recover events, we relied on both videos and timestamps using the subject eye tracker and the handheld HD camcorder. To ensure accuracy a trained member would visually inspect both video types for each subject and each environment watching frame by frame in order to obtain the most precise timing possible. Events were encoded in a spreadsheet correctly categorized and labelled with corresponding timestamps. Every session was reviewed a second time by a different staff member to ensure agreement of both event label and timing. If any disagreements occurred, a third, senior staff member would resolve the issue.

Once all events were recovered and reviewed, the timestamps of each event were encoded into the corresponding EEG file and stored in the 'EEG.event' struct to indicate the event name, category, and timing.

Generally, events to be tagged were categorized by priority as follows:

1) <u>Critical</u>

1A: Conditions (Baseline, Non-stress, Stress)

1B: Button presses (The HED tags differentiate correct from incorrect button presses)

1C: Flag looks (If a subject looked at a flag multiple times, each look was tagged)

1D: Audible feedback (e.g. siren to indicate missed flag)

1E: Voluntary subject actions related to the experimental procedure (e.g. subject opens mouth to receive the swab for the saliva sample)

1F: Technical errors (e.g. video projector playback skips during indoor trial)

2) <u>Likely to be useful for analyses</u>

2A: Terrain changes (e.g. downhill/uphill/flat, dirt, grass, rock, wood chips, boardwalk)

2B: Experimenter instructions (sound played to instruct subject to slow down or speed up their pace)

2C: Involuntary subject actions related to walking (e.g. subject stumbles over a tree root)

2D: Voluntary subject actions that are incidental, related to walking (e.g. subject walks around a mud puddle)

3) <u>Not as likely to be useful for analyses</u>

3A: Distractors (e.g., pedestrian runs down the trail, construction noise is heard)

3B: Involuntary subject actions not related to walking (e.g. subject sneezes)

3C: Voluntary subject actions that are incidental, not related to walking (e.g. subject lifts arm, subject scratches face, subject reads a road sign, subject drinks water, subject whistles)

Events hierarchy used primary and secondary categories shown in Table 1.

Table 1
Commonly used primary and secondary category types in HED tagging along with the most common examples for each.

Primary category	Secondary category	Event
1) Environmental Event	a) distracter	• Pedestrian passes • Helicopter overhead • Truck passes • Animal crosses path
	b) terrain	• Downhill/uphill portion • Puddle in path • Subject trips on branch
2) Scenario Event	a) verbal cue	• Condition change [baseline/stress/non-stress] • Directions • Instructions • Slow down/ speed up
	b) non-verbal cue	• Horn (stressor) • Slow down/speed up • Saliva swab
3) Behavioural Event	a) purposeful	• Saliva swab • Subject speaks
	b) accidental	• Subject trips on branch • Sneeze • Cough
4) Technical Error	a) distracter	• Eye-tracker slipping • Experimenter next to subject • Video skips
	b) non-distracter	• Heart rate monitor falls off ear • Trigger cord disconnects • EEG software stops recording

4.11.3. EEG data

EEG was constantly monitored in real-time throughout the duration of the experiment. Any technical problems with individual channels were noted in a log. If a channel appeared noisy during testing an experimenter applied more gel without interrupting the subject when possible. Fig. 4A (left=indoors, right=outdoors) shows an example of 5 s of EEG data from a single representative subject during baseline walking for 7 common EEG channels across the head (FPz, Fz, Cz, Pz, Oz, C3, and C4).

4.11.4. Inertial measurements units (IMUs) and ground reaction forces (GRF)

IMUs were placed in 6 locations (one on the top of the toe box of each foot, one around the ankle of each foot, one around the chest, and one around the waist). We found that the IMU on the top of the toe box provided the most reliable data for recovering gait events. In order to validate this we matched corresponding peaks in acceleration with known heel-strikes and toe-offs from the ground reaction forces of each force plate. Fig. 4B shows the accelerations of the foot IMUs in the vertical direction (z-axis) aligned with force plate measurements (Fig. 4C).

4.11.5. Cortisol

All samples were sent to a lab on campus using Salimetrics Assay kits. Results were given in concentration (µg/L).

4.11.6. Eye tracking

Eye tracking videos were time synchronized with EEG data and aligned accordingly. Timestamps of eye tracking events were converted to match EEG sampling rate and encoded in the EEG.event struct as described in HED Tagging section.

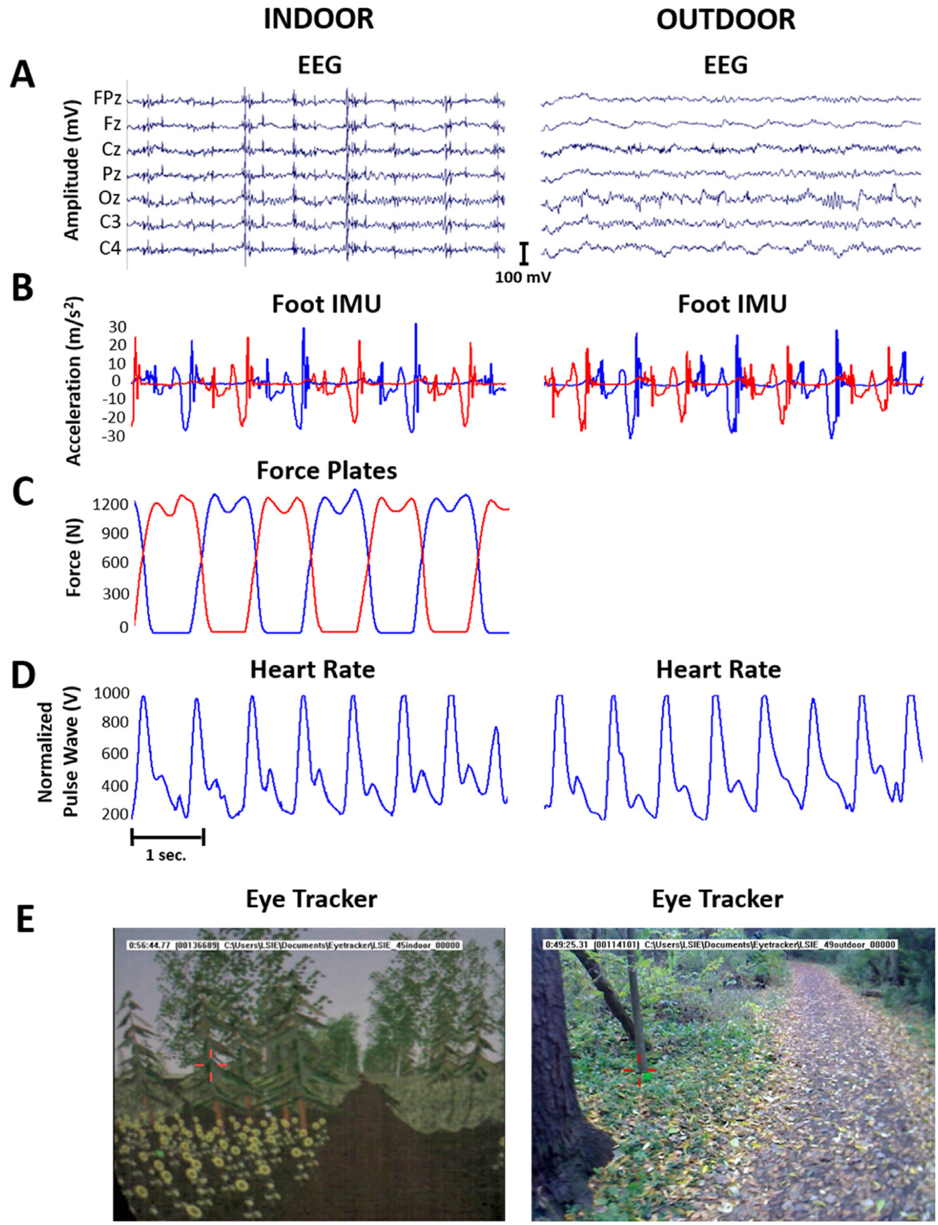

Fig. 4. Example data from a single representative subject. Left column shows example indoor data and right column shows example outdoor data (5 s each) for (A) EEG, (B) both feet IMU z-axis accelerations (blue = left foot, red = right foot), (C) ground reaction forces of both force plates (blue = left foot, red = right foot), (D) pulse sensor heart rate, and (E) screenshot of subject's eye tracker video with crosshair in red indicating eye gaze. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.11.7. Heart rate

Pulse Sensor was linked to an Arduino logging data in real-time. The pulse sensor uses Pulse Transit Time (PTT), which is a measurement of the time it takes for the heart pulse wave to travel throughout your body. In application this measures the time between the R wave of an Electrocardiogram (ECG) against the pulse wave recorded with a Photoplethysmogram (PPG). The Pulse Sensor Amped software amplifies the raw signal of the previous Pulse Sensor, and normalizes the pulse wave around V/2 (midpoint in voltage). The Pulse Sensor responds to relative changes in light intensity. If the amount of light incident on the sensor remains constant, the signal value will remain at (or close to) 512, the midpoint of Arduino's range. For some participants, the Pulse Sensor data was not reliable and didn't provide reliable peak detection or showed clipping. In those cases, we recovered ECG heart rate peaks from EEG electrodes attached to the back of the neck.

Limitations

The goal of the study was to capture an outdoor recording session with as close to "real-world" scenarios as possible that would have data similar to the controlled indoor laboratory recording session. As such, we specifically wanted the outdoor walking speed to be natural for each participant, but needed to keep a steady pace that wasn't too different from the indoor treadmill pace. We did not measure the walking speed during the outdoor condition. The experimenter cue adjustments to walking speed are included in the Hierarchical Event Descriptor (HED) tags for each session however. The outdoor recording session had variations in temperature, lighting, ambient noise, humidity, etc. These factors are all real world parameters that are not controlled in the data set. This is a limitation of collecting data in outdoors real world conditions. There are notes about the weather conditions for each participant in the subject testing log.

Another limitation of real world outdoor conditions is that the surface terrain changed in composition and incline/decline during the recording session. Most of the outdoor walking path was compacted dirt with occasional rocks and pebbles. There are HED tags for major changes in terrain when they occur. We also included both subject eye tracker video as well as "behind the subject" camcorder video to allow for other users to analyse subjective changes in terrain themselves as well as any other features of interest that have not already been tagged. The intent of this manuscript is to provide as complete of a data set as possible to enable other parties to utilize this data for analyses outside the scope of our goals. While we made a great effort to provide as detailed events as possible, there will always be potential features of interest that are missing.

One of the major limitations with utilizing video recordings to tag and analyse events with EEG is the timing synchronization and inaccuracies that may occur. While the EEG recordings used a high sampling rate at 512 Hz, the video recordings were limited to a sampling rate at 30 Hz. For analyses requiring high precision timing accuracy such as event-related potentials (ERPs), the components of interest may require precision within 5–10 msec accuracy while the difference of 1 video frame offset is 33.33 msec. We tried to minimize any errors as best as possible by requiring multiple confirmations of all video event timings by multiple reviewers. We are confident that the synchronization was double-checked for all events and data streams, but it is a limitation of the data set. Our research staff randomly sampled and checked the synchronization of data at multiple stages of procession. We are confident that label timings are accurate to the human eye within 1 video frame. The sampling rate limitations of video at 30 Hz should be taken into consideration when conducting any analyses relying on more precise timing synchronizations.

The flag placement outdoors was chosen by research staff and marked on a map so that it was reproduced for all participants. In general, flags were left in the arboretum throughout testing. Research staff checked each flag and replaced any flags that had been removed prior to each participant testing session. There was also the limitation that the indoor virtual environment was not an exact duplicate of the outdoor environment. The research staff attempted to place the flags in comparable distances and heights across conditions but as the environments were not identical, there is some variation in locations across conditions.

We did not include subjective stress surveys in the data collection. The monetary penalties and auditory alarms were intended to induce changes in stress. Objective analysis of changes in stress would have to be based on the quantitative cortisol concentrations for each condition of the task both indoors and outdoors. The lack of subjective stress surveys is a limitation of the data set.

In complex, real world experiments, it is often possible to have missing or noisy data. A limitation with the current data set is there is noise included in the data as we have chosen to provide raw data with as much details about events and precise timings as possible. Rather than cleaning the data and providing highly processed data, the raw data allows researchers to choose and test their own cleaning algorithms. The only data that was noticeably missing was Pulse Sensor data on heart rate, as it was not reliable and didn't provide reliable peak detection. As a result, we recovered ECG heart rate peaks from EEG electrodes and included that in the data set to analyse heart rate.

As with any complex real world data set, there is a limit about how much information and data specifics can be anticipated by new users. A limitation of this data set is that it does not contact a tutorial to teach you how to analyse the data. Future questions about the data can be directed to the authors at Grant Hanada (ghanada510@gmail.com) and Daniel Ferris (ferrisdp@gmail.com).

Ethics Statement

All research was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki). Our protocol was approved by the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board for the protection of human participants (protocol number: HUM00066867). All participants provided written informed consent before participating.

Credit Author Statement

Grant Hanada: Conceptualization, Methodology, Software, Validation, Data curation, Visualization, Writing –original draft, Writing –review & editing; **Marija Kalabic:** Methodology, Software, Validation, Data curation, Visualization, Writing –review & editing; **Daniel P. Ferris:** Conceptualization, Writing –review & editing.

Data Availability

LSIE individual subjects full set (Original data) (figshare)

MOBILE BRAIN-BODY IMAGING DURING INDOOR TREADMILL WALKING AND OUTDOOR OVERGROUND WALKING WITH A VISUAL SEARCH TASK (Original data) (IEEE Dataport)

Acknowledgements

This research was supported by ARL W911NF-10-2-0022 and NSF BCS-1835317. The authors would like to thank Bryan Schlink for helping with experiment design, technical expertise, and data collections. We thank Michael McBurney for his help reviewing and validating the data. We thank Andrew Moseley-Gholl for his supervision and organization of our staff. We thank Ramses Alcaide for helping design and develop the initial ideas for this project. We also thank all our undergraduate staff including Renee Philson, Kirk Detloff, Sandy Mouch, Maggie Smith, Noel Deshpande, and Rayen Patnaik for their tireless efforts in data collection and review.

Declaration of Competing Interest

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