D3.8 – Annotated Open datasets

Project No	GA824160
Project Acronym	EnTimeMent
Project full title	ENtrainment & synchronization at multiple TIME scales in the MENTal foundations of expressive gesture
Instrument	FET Proactive
Type of action	RIA
Start Date of project	1 January 2019
Duration	48 months

Distribution level	[PU][1]
Due date of deliverable	Month 36
Actual submission date	31 January 2022
Deliverable number	D3.8
Deliverable title	Annotated Open datasets
Type	ORDP (Open Research Data Pilot)
Status & version	Final
Number of pages	40
WP contributing to the deliverable	3
WP / Task responsible	UCL
Other contributors	ALL
Author(s)	UCL
EC Project Officer	Christiane Wilzeck
Keywords	**Dataset, body movement, multimodal sensors**

[1] **PU** = Public, **PP** = Restricted to other programme participants (including the Commission Services), **RE** = Restricted to a group specified by the consortium (including the Commission Services), **CO** = Confidential, only for members of the consortium (including the Commission Services).

Contents

1 Introduction 6

2 Scoping review of open datasets on movement [Olugbade et al., in submission] 6

2.1 Systematic survey process 7

3 EmoPain@Home: EnTimeMent movement dataset of people with chronic pain in their home (UCL) 12

3.1 Sensors setting 12

3.2 Participants 12

3.3 Data collection protocol 14

4 Description of the Genoa full body action and emotion set and validation data (Poyo Solanas & de Gelder) 17

4.1 Stimulus set description 17

4.2 Video processing 18

4.3 Behavioural experiment 21

5 EMOSYNC 21

5.1 Description 21

6 Spook and Play 24

7 Hindustani raga and singer classification using pose estimation *(Durham)* 28

7.1 Description 28

7.1.1 Brief description 28

5.2 Parameters and initial values used for analysis 30

5.3 Licences 30

5.4 Ethical approval 30

5.5 References 31

7.1.2 *Publications employing this corpus* 31

7.1.3 *Pose estimation algorithms* 31

8 Qualisys markerless mocap campaign – November 2021 31

9 Methodology and dataset description for the project entitled ' Individual limb-dependency in a curvilinear rhythmic movement' (Zarandi, Z., Stucchi, N., Fadiga, L., Pozzo, T.) 33

7.1 Participants 33

7.3 Measures 34

7.4 Extracted features 35

10 UCLIC-Runner dataset (Bi, T., Berthouze, N., Singh, A., & Costanza, E. (2019)) 36

11 References 38

Abbreviations

EU	European Union
EC	European Commission
WP	Work Package

1 Introduction

This deliverable provides an overview of the datasets collected within this project and that will be made available to the community. Section 2 presents a systematic review of movement datasets currently available for use within the research community. The rest of the deliverable presents brief descriptions of datasets being created within the EnTimeMent consortium. Each of these datasets will be open to the community as soon as related work has been published. See Table 1 for a quick overview of these EnTimeMent open labelled datasets. Descriptions of datasets not yet described in other deliverables will follow after Table 1. Datasets created during the Markerless Qualisys campaign (November 2021) to evaluate the use of markerless motion capture system in the EnTimeMent scenarios are currently being preprocessed at Qualisys. These datasets will be added to the table once the data are ready.

2 Scoping review of open datasets on movement [Olugbade et al., in submission]

While the 16 existing movement dataset reviews we found have their merit, they are limited in their coverage in terms of size and research discipline. It is critical to additionally have a comprehensive overview that is based on a systematic survey and cuts across disciplines to enable higher visibility of available datasets to the research community and further foster interdisciplinary collaborations. As such a systematic survey of open movement datasets was carried out by UCL, and these were then critiqued by the consortium partners to understand: 1) the limitations of the current datasets in supporting movement research, and 2) the factors that may facilitate sharing of datasets within and across disciplines.

The survey examined over 80,000 publications and retrieved more than 1,500 datasets (including those not open for secondary use) from them; about 30 additional datasets were added from previous knowledge. Based on these, the collaborative paper presents to the research community a total of the 704 open datasets. We describe these datasets are based on 10 different variables that can be valuable to researchers searching for secondary data. These variables include: name, purpose, type of data (sensors, type of movement and setting), annotation types, data source, population group, number of people simultaneously recorded, type of mocap if used, funding source. We additionally provide an analysis of these datasets and further review them under the themes of diversity, ecological validity, and the data recorded. Our discussion resulted in an 11-dimension framework that can be used as a tool to guide researchers in planning the creation of open movement datasets. Ethical and data protection issues are also included in this framework given their critical importance in collecting and sharing datasets. Finally, we call for the development of an open-access repository for human movement datasets and related resources, which we argue will be valuable in advancing work in areas of movement research.

This work has been the interdisciplinary effort of the EnTimeMent researchers across the consortium from multiple disciplines including affective computing, clinical psychology, disability innovation, human-computer interaction, machine learning, music cognition, music computing, and neuroscience.

Details on the survey, results, and discussions are in [Olugbade et al., in submission]. The publication and dataset catalogue will be available when the paper is accepted. The systematic survey process is described below as proof of soundness.

2.1 Systematic survey process

The first author (T. Olugbade) conducted a systematic search of relevant articles via 3 search engines for scholarly literature (Google Scholar, PubMed, APA PsychInfo) and 3 publisher repositories (ACM Digital Library, Springer, IEEE Xplore Digital Library) between 21 October 2020 and 28 January 2021. These databases were carefully selected for their comprehensive coverage of peer-reviewed research or other technical articles in the pertinent areas of human science and computing. For each database, we tailored our search according to the search functionalities available for the database; but in general, we searched for articles that described human movement data. Key search terms (reported in the paper) were discussed with a multidisciplinary team of researchers.

The first author followed two levels of shortlisting to obtain the final set of datasets. For the first level, they went through all titles and abstracts of the search results and excluded duplicate description of the same dataset, books/theses, patents, citations without full-text available, survey papers, or descriptions of non-human movement data. This resulted in 4,663 relevant articles. For the second level, they went through the full text of each of these articles and further excluded articles that were found to meet the exclusion criteria above, could not be accessed, were not available in English language, described datasets based on a single anatomical region (e.g., face only), presented simulated data, described still images or far/top view videos, or unusably limited description of the movement data. Of the remaining articles, 1,599 of them were found to be secondary references for the datasets that they described. For each of these, the first author searched for the primary article or website.

A final list of 1,692 datasets (with 278 based on the secondary references found in our systematic search and 34 found completely outside of our systematic search, e.g., from a priori knowledge) was obtained. A charting of these datasets was done across 10 variables described in the publication in submission. The catalogue, analysis, and review in the publication focused on the 704 datasets that we could ascertain are open to the research community for secondary use. To check for data availability for use by the research community, the first author carefully did a manual search of the abstract, conclusion, dataset description sections, and footnotes of the associated article. They additionally did an automatic search for relevant keywords in the text of the article, particularly ‘available’, ‘access’, and ‘obtain’. For named datasets, they further searched the Internet for websites or other resources with information about how the dataset could be accessed.

Table 1: quick overview of @EnTimeMent open labelled datasets

Name	Motivation	N. of people	Context	Sensor data	Activity	Label	Other data	Ethics	Description & publication state
EmoPain@Home	Developing protective behaviour and pain level detection in people with chronic pain	9 people with chronic pain over multiple days and activities	People’s homes	Notch motion capture (6 IMU sensors)	A variety of exercise and functional activities	Pain level + participant condition	Interviews & diaries	Consent to share motion capture data	D3.8 – Not open yet
UCLIC-Runner dataset	Developing affect and pain detection in runners	11 runners	Outdoor spaces of runner’s choice	Insole pressure sensors, Empatica physiological & wrist movement sensors, head motion sensor	Running	Pain, exertion, motivation, affective valence levels	Interviews & diaries	Consent to share all data sensors	D4.7 – Not open yet
Dataset Genoa	Investigating midlevel body features	19 people	Empty controlled	Pose estimation data derived from OpenPose	5 emotional actions, 5 neutral actions	Emotion, action	Behavioural ratings on the videos	Consent to re-use for research	Not open yet Pending publication

	and their relation to emotion and action recognition.		(green) background)						
EMOSYNC	Testing effect of emotion on the improvisation dance in a triad	13 triads	Dancing in a MoCap lab	VIcon MoCap Trigno ECG	Dancing with one arm, improvisation		Psychometric information, mood scale, empathy	Anonymised ok to share	Not open yet Pending publication
Spook and Play **CHRONOS_IADS**	Testing effect of emotion on the ability to synchronise with others	60 people, 15 groups of 4	Playing Chronos game (disembodied sync)	LEAP sensors	Unilateral hand synchronisation task		Psychometric information, mood scale, empathy	Anonymised ok to share	Not open yet Pending publication
Hindustani raga and singer classification using pose estimation	Exploring movement of singers and small ensembles while performing	3 singers (solo recordings),1 singer with tabla accompaniment (duo recordings), samples of full ensemble	Lab/ concert stage	Pose estimation data published with video data	Music performance	Music item, singer		Consent to re-use for research	Clayton, M., Li, J., Clarke, A. R., Weinzierl, M., Leante, L., & Tarsitani, S. (2021, October 21). Hindustani raga and singer

		concert performances							classification using pose estimation. https://doi.org/10.17605/OSF.IO/T5BWA
Individual limb-dependency in a curvilinear rhythmic movement	Exploring individual motor signatures in upper arm motor control	40 healthy participants	Lab	Wacom tablet	Individual motor control	Action		Anonymised ok to share	Not open yet Pending publication

3 EmoPain@Home: EnTimeMent movement dataset of people with chronic pain in their home (UCL)

EmoPain@Home is the first dataset containing movement of people with chronic pain engaged in activities within their home. These activities include both functional activities as well as exercises, and they were chosen by the participants themselves according to the types of activity they do in their home but that they often struggle with. Hence activities differ between people. We maintained the definitions participants gave to each activity.

Ethical approval was obtained by the UCL REC Project Ref: 5095/001. Participants provided written consent for the sharing of the motion capture data and their labels.

3.1 Sensors setting

Participants' movements were captured using low-cost movement sensors (Notch: https://wearnotch.com/). These were placed on the trunk, hips, and on the limbs on one side of the body (e.g., right arm and leg), for a total of 6 IMU sensors. The data were tracked by the Notch app running on a smartphone. Rather than using full body mocap suits, we aimed to minimize the number of sensors to simulate possible future use of such devices. The position of the sensors was based on previous studies aimed at automatically predicting the level of pain perceived by the person from their movement (Olugbade et al., 2018). The study showed that the data from one side of the body was generally sufficient. The decision to use a minimal set of sensors was threefold: 1. reduce the cost of the wearable device; 2. reduce interference between sensors; 3. reduce the computational cost to allow software to run locally on the phone and thus minimise privacy concerns.

3.2 Participants

Participants were recruited from community chronic pain groups. 10 participants were recruited, but one withdrew due to personal circumstances. The participants comprised 5 female and 5 male between the ages of 27 and 59 (1 participant in their 20s, 2 participants in their 30s, 3 participants in their 40s, and 4 participants in their 50s; mean=45.11, standard deviation=11.50). All participants recruited self-identified as living with musculoskeletal chronic pain (CP) involving the lower back area, although the severity of their condition greatly varied. 4 of the participants reported sciatica as their CP condition; 2 participants reported CP resulting from an old spinal injury; and the 3 others reported other CP.

Table2. Types of activities recorded. The * on the ID means that the person took part in both supervised and unsupervised sessions. Activities in bold were only carried out in unsupervised sessions, whereas activities underlined were carried out in both supervised and unsupervised contexts.

<table>
<tr><td>ID</td><td>Challenging activities</td><td>Not challenging activities</td></tr>
<tr><td>P002*</td><td><u>Changing bed sheets</u>
<u>Washing up</u>
Loading washing machine
Unloading washing machine
Clean windows</td><td>Power walking
Sweeping floors
<u>Dusting</u>
Hoovering</td></tr>
<tr><td>P003*</td><td>Hoovering
<u>Washing up</u>
Clean bathroom
Unloading shopping
Cleaning windows
Tidying up</td><td><u>Washing up</u>
<u>Unloading washing machine</u>
<u>Loading washing machine</u></td></tr>
<tr><td>P004*</td><td><u>Hoovering</u>
<u>Changing bed sheets</u>
Vacuuming inside of car
Watering garden</td><td>Clean bathroom
Dusting inside of car
Preparing food
Cleaning parrot cage</td></tr>
<tr><td>P005</td><td>Painting shelves
Painting wall
Power walking
Cleaning bathroom</td><td>-</td></tr>
<tr><td>P006</td><td>Changing bed sheets</td><td>Unload dishwasher
Organising boxes</td></tr>
</table>

P007	Unloading washing machine Unloading dishwasher Changing bed sheets	Loading washing machine Loading dishwasher Tidying up
P008	Washing up Hanging clothes to dry Hoovering Changing bed sheets Cleaning windows	Yoga Unloading washing machine
P009	Ironing Preparing food	Filing documents
P010	Tidying up	-

3.3 Data collection protocol

The full study involved 3 one hour-long sessions, and a half-hour training session. All study sessions were carried out with the researcher present. The entire study was conducted remotely and in compliance with Covid-19 health and safety protocols. Participants were sent all the equipment needed via courier; interviews and observations were conducted via Microsoft Teams. The researcher facilitated the use of the sensors and of the app via Microsoft Teams.

3.3.1 Data collection activity

For each participant, the data collection took part over three days within a week or so. This was to maximize the amount of data collected from the same participant, and also to maximize variability in pain level during the data collection.

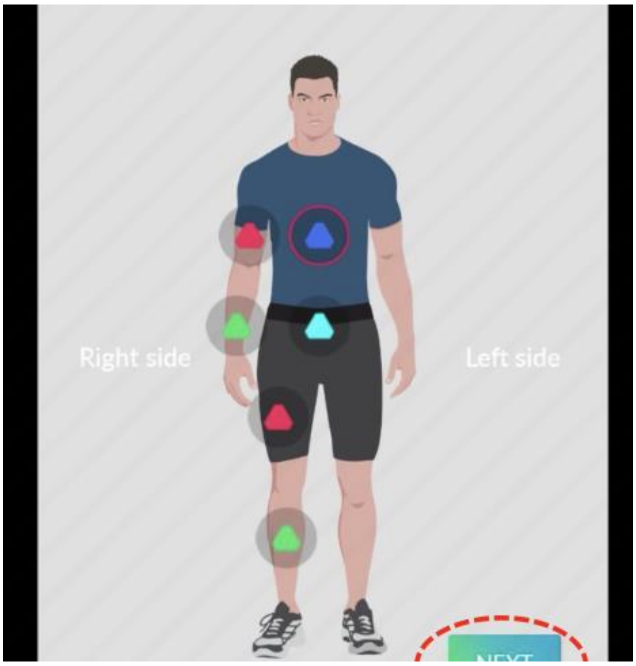
Participants were asked to perform indoor household activities that they would normally carry out as part of their daily routines. These activities were agreed with each individual participant in advance of data collection during a pre-diary study and a short interview.

Participants performed between 1 and 3 activities each session, with each activity lasting between 5 and 30 minutes (average 15 minutes). Activities included chores as varied as cleaning surfaces, hoovering floors, painting, washing dishes, loading/unloading washing machines, etc. Before starting each activity, participants were asked to label it as either 'challenging' or 'not challenging'.

While carrying out each activity, participants' movements were recorded via video and motion capture sensors. Each participant wore a set of 6 Notch sensors at the following positions (see Figure 1): 1 x right wrist; 1 x right upper arm; 1 x chest; 1 x waist; 1 x right thigh; 1 x right calf. In some instances, one of the motion capture pods failed to turn on: in that case, participants were asked to record only using 5 sensors (removing the calf one).

Some participants repeated the data collection on their own (without the remote presence of the researcher). In this case they may have repeated the same activities decided in advance, or engaged in others also considered challenging, but in either case without a video recording.

Figure 1: position of sensors



3.3.2 Self-report

In order to label the sensor data, participants were also asked to self-report each minute while performing activities. In particular, they were instructed to respond to the following prompt: "What is your level of pain, worry, and confidence at this moment in time?".

Pain and worry were assessed on an ascending numerical scale from 0 to 10, where 0 equalled to ‘no pain/worry’ and 10 to ‘very severe pain/worry’; confidence was assessed using an ordinal scale of {no confidence; less than average confidence; average confidence; more than average confidence; maximum confidence}. In order to avoid disruption to the overall flow of the physical activity, the long-form question above was replaced by a shorter prompt (i.e., *‘time’*)

3.3.3 Diary study and Interviews

Before and after the data collection, participants were interviewed to identify activities of interest for the participant and prepare for data collection in the home. It also enabled the researcher to get a view of the home (remotely) and consider the positioning of the participant’s laptop to follow and record the activity. An interview at the end of each session aimed to better understand the use of wearable technology to support chronic pain management. A more detailed description of the full study and results from the interview is provided in D4.5.

Additionally, participants were asked to fill in a diary twice a day while taking part in the study. The diary served a dual purpose: it was a reflection tool for preliminary thoughts about some of the topics discussed during interviews, and it enabled the researcher to get a sense of how participants’ reports of pain, confidence, worries, plans, etc. varied through the day, and between supervised moments (i.e., interviews), and unsupervised ones (i.e., diary).

Figure 2. Still from mocap output of movement

4 Description of the Genoa full body action and emotion set and validation data (Poyo Solanas & de Gelder)

4.1 Stimulus set description

Nineteen actors (six excluded due to invalid or incomplete recordings; all Caucasian, mean ± standard deviation (SD) age=27±5,52 years; range age= 21-39 years; seven males) were instructed to express five neutral (N1-N5) and five emotional actions (E1-E5) for five repetitions (see Table 3). Actors were instructed to dress in black and almost all the actors wore black clothes (16 out of 19 actors), however, the majority of the actors did not wear black shoes (13 out of 19 actors). The videos were recorded (50 frames per second; 1280x720px) using two synchronized Camcorder HCV (JVC, model GY-HD251) at a fixed position recording the actors from a frontal view in a background- and light-controlled environment.

Table 3. Overview of all the categories and corresponding emotion and action labels.

Emotional or neutral category	Category label	Intended emotion	Intended action
Emotional	E1	Fear	Self-protecting
	E2	Happiness	Greeting a friend
	E3	Anger	Raging
	E4	Dominance	Showing dominance
	E5	Submission	Subordination
Neutral	N1	Neutral	Brushing off
	N2e	Neutral	Eating a banana
	N2p	Neutral	Peeling a banana
	N3	Neutral	Eating berries
	N4	Neutral	Searching an object
	N5	Neutral	Catching a ball

4.2 Video processing

4.2.1 Cutting process

The recordings were cut into one-second videos, using Adobe Premiere Pro 2021 (Adobe Systems, 2021). Because some actions lasted longer than one second, some recordings were cut into several one-second videos, which led to the differentiation between two different actions within the action category N2; ‘Eating a banana’ and ‘Peeling a banana’. It was therefore chosen to divide the N2 category into two neutral categories. See Table 3 for the final categories and corresponding intended action and emotion.

4.2.2 Background and face removal

After cutting the videos to one-second duration, the background was replaced by a homogeneous green color (RGB (98, 218, 149)) using the methodology developed by Lin and colleagues (2020). The standard settings were followed, with the exception of the model backbone scale which regulates preciseness of the body outlining, that was set to the maximum value: 0.5. After the background removal, faces in the videos were blurred in order to restrict faces from having an impact on emotional body perception. The face blurring was performed using several algorithms in MATLAB (Matlab vR2019a, The MathWorks Inc., Natick, MA, USA). First, the face position was estimated using a face detection algorithm: the Viola-Jones algorithm (Lienhart, Kuranov & Pisarevsky, 2003; Ojala, Pietikäinen & Mäenpää, 2002; Kruppa, Castrillon-Santana & Schiele, 2003; Castrillón, Déniz, Guerra & Hernández, 2007; Shiqi, n.d.; Viola & Jones, 2001; Dalal & Triggs, 2005). Then, a square was set around the estimated face and the center point was calculated and tracked over all frames of the video with the Kanadeo-Luca-Tomasi algorithm (Lucas & Kanade, 1981; Tomasi & Kanade, 1991; Shi & Tomasi, 1994; Kalal, Mikolajczyk & Matas, 2010). The center point was used as the middle point for the inserted Gaussian filter (size [75,75], σ=100) and the blur circle radius was determined by the face detection algorithm and the coordinates of the center point.

4.2.3 Solving background removal and face blurring problems

After these steps, several problems surfaced related to the background and face removal. Regarding the face blurring, sometimes the face was not correctly recognized by the algorithm, or the center point was lost or misplaced over frames. Since this study aimed at having five videos per actor and category, it was examined if the incorrect videos could be replaced by other videos from the many repetitions within the same recording. For the categories for which this was not possible, the face estimation algorithm was replaced by another algorithm; OpenPose (Cao, Hidalgo, Simon, Wei & Sheikh, 2021; Osokin, 2019). At last, if the latter algorithm was not able to solve the face blurring problems, new video cuts were made in which the actor was looking forward at the beginning and end of the video, to improve recognition of the face by the face detection algorithm. These cuts were sometimes longer than one second. Eventually, all the new cuts were processed by the background removal algorithm and face blurring algorithm, as described above, and were, if necessary, cut into one-second videos.

In addition to the artefacts caused by the face blurring algorithm, the background removal caused artefacts. For example, sometimes a body part was removed or the knees were not separated correctly and merged. Furthermore, sometimes the ground under the feet was not completely removed and flickered over frames. To solve these issues, incorrect videos were replaced by correct videos. For the cases in which this was not possible, the model backbone scale was lowered to 0.45, 0.35, or 0.25, until the artefacts were disappeared or minimized.

4.2.4 Stimuli pre-selection

Because many categories consisted of more than five correct videos after all processing steps, an experiment was performed to select the best five recognized videos. Each video was rated by three different lab members (N=8). The five best stimuli were chosen based on rating percentages and the absence of artefacts. If a video with artefacts was rated among the best five, it was replaced by another randomly chosen one without artefacts. If there was an equal rating percentage, the video with the least editing artefacts was chosen. If not applicable, a random video was chosen. Finally, the total stimulus set consisted of 1045 videos of which 570 representing neutral actions (19 actors x 5 repetitions x 6 categories; category N1-N5) and 475 representing emotional actions (19 actors x 5 repetitions x 5 categories; category E1-E5), see Table 3.

Given that the project focuses on identifying the features that differentiate between actions and therefore a certain level of homogeneity of the exemplars of each category is important. Some categories displayed large variability and this could be reduced by using mirror version of some videos, for example when there was large variability in the hand used or in the left-right direction of the action. Furthermore, in the category 'searching an object' (N4), the majority of actions represented squatted down actors and a minority represents a downward movement from standing to squatting. For the same reason as many videos as possible were chosen in which actors were squatted down.

4.3 Behavioural experiment

A behavioural experiment was performed in which participants rated the current videos regarding mid-level features and their cognitive interpretations of the actions. The questions regarding mid-level features (body contraction, amount of overall and vertical movement, movement direction, and head inclination) had to be answered on a five-point Likert scale. Furthermore, to gather information about the subjective cognitive interpretation of the new stimuli, seven more questions were asked about high-level features (emotion and action categorization, body part attention, valence, arousal, familiarity, and naturalness). The questions about emotion and action categorization, and attended body part were multiple forced-choice questions and all other questions were asked on a five-point Likert scale.

In total, 81 participants took part in the study (59 Female; mean ±SD age=22±2,67; range age=19-31; Nationality= Dutch: N=42, German: N=15, European (non-Dutch or German): N=15, Eastern: N=6, American: N=2, unknown: N=1, see appendix B for more details). All participants had normal or corrected-to-normal vision and no medical history including neurological or psychiatric disorders. Participants were recruited via SONA, a research participation system of Maastricht University, or via social media. Participants received SONA credits, a voucher, or monetary compensation for their participation. This research was approved by the ethical committee of Maastricht University and was performed in accordance with the declaration of Helsinki. All participants had signed informed consent prior to the commencement of the experiment.

5 EMOSYNC

5.1 Description

EMOSYNC experiment was conducted on right-handed or ambidextrous participants that reported enjoying dancing. Groups of three participants without prior acquaintances, paired by same sex were randomly generated. According to Issartel, Gueugnon, and Marin (2017), the participants were controlled in their levels of expertise in improvisational type of tasks into novices, intermediates and experts. There were 11 groups of novices and two groups of intermediate dancers.

The task was to produce complex, varied and interesting movements with their right hand. No explicit instruction to synchronize was given to the participants. The duration of each trial was 30 seconds. Three affective states were induced: positive emotion, neutral state and negative emotion. There were 18 solo trials and 15 groups trials per experiment. Each participant performed 6 solo trials, 2 trials per emotion, and 15 group trials, 5 trials per emotion. The emotional induction was carried out by a manipulated feedback type of task and we used the interactive platform Mentimeter to implement it. Essentially, the participants needed to rank each other with respect to their perception of the task success and were given the manipulated feedback.

Ethical approval was granted by a local ethical board (IRB EuroMov 2002A) in accordance to the Declaration of Helsinki. Each of the participant followed strict social distancing protocol for COVID-19 prevention. There were three modalities recorded during the experiment: behavioral modality, psychological and physiological. For behavior modality Vicon Motion Capture cameras were used. Camera calibration were performed daily before the start of experiment. The shutters were closed so that the light would not interfere with the signal. The frequency of the recording was 100 Hz. There were 9 retro reflective motion capture markers. Seven retro reflective markers were placed on the right hand and two were placed on the head. Specifically, the markers were placed on the following anatomical landmarks: RSHO - the right shoulder at the acromio-clavicular joint, RUPA - the right upper arm between shoulder and elbow, RELB - the lateral epicondyle of the right elbow, RFRA - the right forearm between the elbow and wrist, RWRA - the medial right wrist, RWRB - the lateral right wrist, RFIN - the right hand second metacarpal head, RFHD - the right forehead, LFHD - the left forehead. The obtained signal was position in x, y and z directions.

For psychological modality several measures were carried out. On the one hand, we asked participants to fulfill two psychological measures that could explain the potential effect of emotion before the experiment started: Positive and Negative Affective States (PANAS) (Gaudreau, Sanchez, & Blondin, 2006) which is the French adaptation of Watson, Clark, and Tellegen's work (1988) and The Big Five Inventory français (BFI-Fr) which is the French validation of the BFI in English (Goldberg, 1993). Also, we obtained scores from Emotional Contagion Scale (Raymond, 2017), which is the French adaptation of the original questionnaire published by Doherty (1997), right before the end of experiment. It was done at the end and not at the beginning with other questionnaires so to avoid giving potential hints about the real goal of the experiment, i.e., to uncover the effect of emotion. On the other hand, we measured the psychological dynamics caused by the emotional induction by using the valence and arousal dimensions from the Self-Assessment Manikin (SAM) questionnaire (Bradley & Lang, 1994) as well as resulting social consequences of the activity in the form of 1-item scale of whether a given participant would desire to continue performing with the members of its group. SAM dimensions and social consequence item were measured after each trial.

For physiological modality, we used Delsys Trigno ECG biofeedback sensor with a recording frequency of 2148 Hz. The placement of sensors is a 3-lead ECG. The extracted signal is a one-dimensional vector measured in mV. The recording started at the beginning of the experiment and it continued throughout the full duration of the experiment. Additionally, we used one analog Delsys sensor to link ECG with the Vicon recordings. Specifically, each time when the Vicon recording was launched, this caused an activation in the analog sensor which was later used as an indication of the start of trial.

Table 4. Summary of contents. The collection comprises the following:

13 files for the ECG recordings, one for each triad. CSV format.
429 recordings for the Vicon motion capture recording (head and arm position), one for each trial in a given triad; CSV format.
Anonymised Emotional Contagion scores, matched to the participant by unique ID. CSV format.
Anonymised SAM ratings of the stimuli received for each trial, Valence and Arousal ratings. CVS format.
Anonymised PANAS questionnaire scores. CSV format.

Anonymised Mentimeter responses. CSV format.
GitHub link for the pipelines: https://github.com/EnTimeMent/EMOSYNC_EyeGaze; https://github.com/EnTimeMent/EMOSYNC_Behavioural_sync

6 Spook and Play

6.1 Description

In this study we aimed to investigate the linkage between the emotion and human performance in the digital, disembodied mirror game for small group (n=4). We have tested 15 groups in this set up collecting data from 60 participants. To induce emotion in participants during the experiment we have used acoustic stimuli from the IADS-2 (Bradley & Lang, 2007) stimulation elicited appropriate perception of the emotional valence and arousal in tested sample. Sound is a portable and powerful source acoustic, emotion-laden information, that can be transmitted while participants are attending to another visual stimuli/task. In a study by Nardelli et al. (2015) demonstrated that participants (female) reacted to IADS-2 stimuli with changes in their heart rate (HRV) variability to the valence and arousal in a consistent pattern. In another study Martin-Soelch (2005), those HRV reactions were not found to be susceptible to habituation in a week interval study. Martin-Soelch (2005) investigated the skin conductance and zygomatic muscle activity as well, showing a consistent pattern of psychophysiological responses for IADS-2 stimuli, validating the perceptual self-report data from Technical Report. Unlike in aforementioned psychoperceptual studies that used IADS-2 battery, here participants were engaged in a motor task of group synchronization. Since the problem with induction of the emotion in the lab setting have been raised in the past by the body of research, we believe that is it an important example to show possibilities of bringing this set up to the movement-oriented studies. Acoustic stimuli unlike the visual information does not consume the same attention resources and can be delivered to each individual in a discrete fashion (via headphones) which is advantageous feature for group scenarios, where multiple agents are concerned.

Participants were recruited for the study by a snowballing strategy from the student and staff pool of STAPS University of Montpellier. Ethical approval was granted by a local ethical board (IRB EuroMov 2005A) in accordance to the Declaration of Helsinki. Each of the participant followed strict social distancing protocol for COVID-19 prevention. 60 participants (28 females) participated and were randomly assigned to 15 groups of 4 people mixed sex and age participants (age M=26+/- 6 years). Since the vision the Chronos display movement of other players is anonymous, there is no need to control for the acquaintance level or participants being from the same age and sex group. All participants were neurologically healthy and had normal or corrected to normal vision and normal hearing, 3 were left-handed.

Chronos is an open-source software that allows for egalitarian inclusion of participants regardless of their gender, age (social cues). This is achieved by movement of all participants being translated in real time from their personal space to gaming display (see Alderiso et al., 2017). The positional data was recorded with the LEAP motion sensors picking up the position of the right index finger of each participant as previously reported in Alderiso et al. (2017), with a frequency of 55Hz. The multiplayer mirror game setup was running in a complete topology without a predesignated leader. The acoustic emotional stimuli were delivered via headphones, with matched volume levels across computers, in each trial (randomized order across participants), with a delay of 6s to 8s post launch of each trial (this was done to avoid habituation of the participants to the occurrence of sound and create and temporal expectation, see Figure 3). White noise was sampled for the remainder of the trial. We have selected 9 negatively valenced sounds, 9 positively valenced sounds and 9 neutral sounds. The average rating of the IADS-2 of the sounds selected was for pleasure (Negative: M=1.85, SD=1.44; Neutral: M=4.82, SD=1.71; Positive: M=7.39, SD=1.74) for arousal (Negative: M=7.6, SD=1.9; Neutral: M=4.82, SD=1.71; Positive: M=6.42, SD=2.04) and dominance (Negative: M=4.79, SD=1.87; Neutral: M=4.56, SD=1.99; Positive: M=5.68, SD=1.85).

Figure 3. Construction of each trial in Spook and Play experiment.

START OF THE TRIAL	6-8 SECOND DELAY	SOUND: 6 SECONDS	TRIAL CONTINUATION	END OF TRIAL
	45 SECONDS			

We have chosen the sounds that do not contain musical structure, rhythm (such a clock ticking) or too much verbal communication (in English, as tested sample was primarily native French speaking). Negative and positive sounds were chosen to proportionally differ as much as possible on the valence spectrum, but were matched for arousal and dominance using values from IADS-2 Technical report (with higher arousal scores being predesignated to negative sounds in this battery as previously reported in the literature).

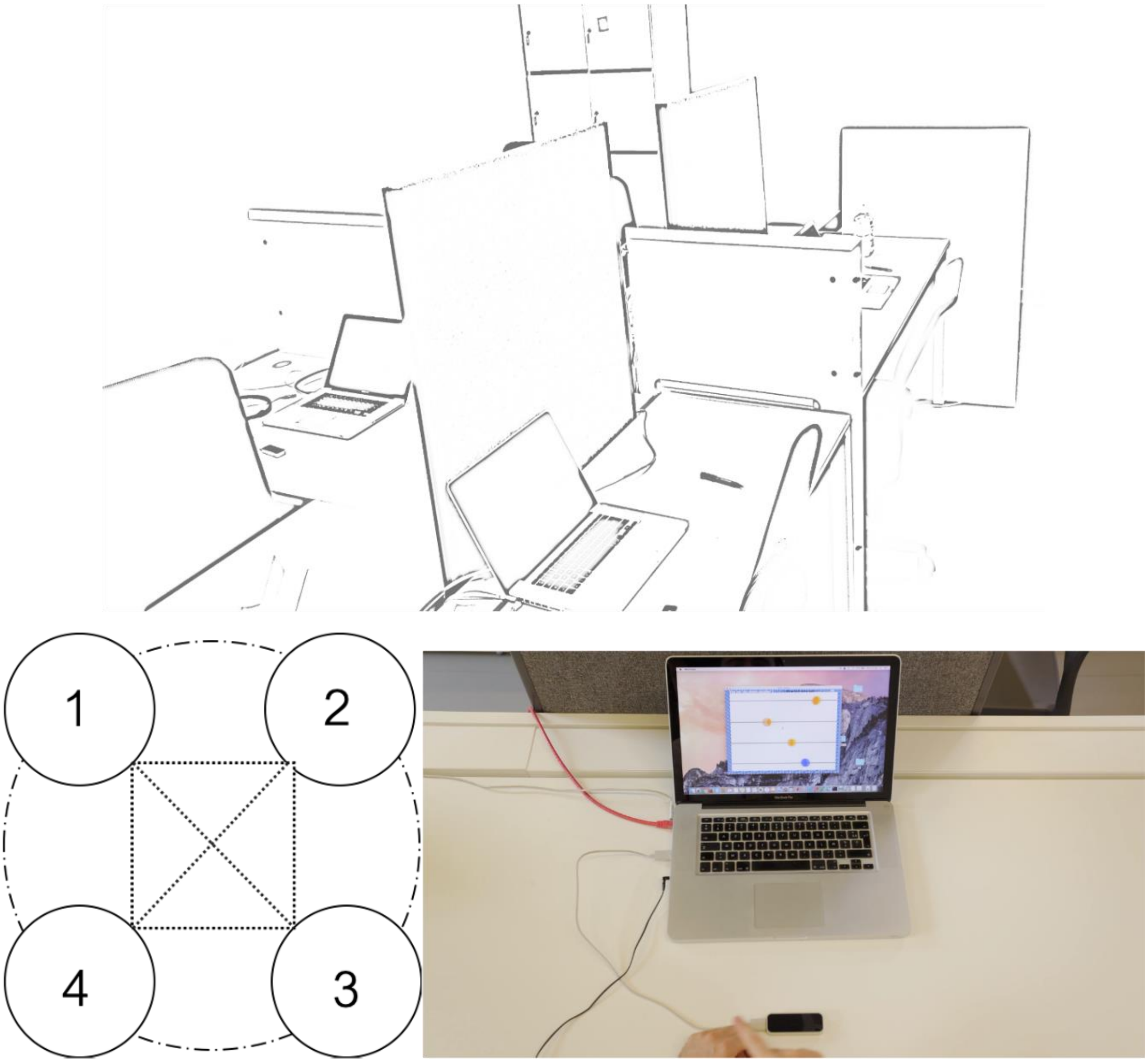


Fig. 4. Laboratory setup - player booths. Top panel depicts the experimental room with the shields between the participants to prevent view and real-time identification of the other members of the group. Left bottom panel shows the complete (Bardy et al., 2020) coupling structure between participants using the digital display of avatars (blue is avatar of the participant – always position on the bottom of the screen, orange balls depict avatars of other members of the group; right bottom panel).

Before the start of the experiment each participant was seated in front of the computer display, separated from other players by physical distance and no direct visual contact (see Fig. 1 for the schema of the set up). The workings of the LEAP motion sensor were explained and demonstrated by the researcher - the required finger movement with right index finger and the optimal capture position and distance. Participants were instructed during the Solo trial to create an interesting non-periodic motion representing their motor signature and natural frequency (Słowiński et al., 2016) with neutral sound being played. Then participants were asked to fill in two short questionnaires - selected items of PANAS (to assess their current emotional state prior to the experiment - Gaudreau et al., 2006 (French validation), and The Emotional Contagion scale (Doherty 1997, adapted by Raymond, 2017) to assess their empathetic abilities.

After Solo trials participants were asked to synchronize their motion with that of the circles shown on their respective computer screen, representing the movements of the other agents topologically connected with them. However, players had no global information of the topology of their interactions with other players. Participants performed 54 trials together, of 45 second duration each (this created multiple and balanced combinations of the neutral, positive and neutral stimuli across the group (4 participants x 18 trials per condition x 3 emotional conditions). After each trial participants were asked to fill on a pencil and paper answer sheet, with number of consecutive trials, their perception of the stimuli displayed during the trial. Two dimensions of pleasure, arousal, and dominance, the Self-Assessment Manikin (SAM), an affective rating system developed by Bradley & Lang (1994). In this system, a graphic figure depicts values along each of the 3 dimensions on a continuous scale indicating emotional reaction to a stimulus presented. We have obtained permission to use SAM scale for the EnTimeMent project from CSEA media (copyright owner).

Table 5. Summary of contents. The collection comprises the following:

Matlab dataset of 60 participants performing the synchronization trials, in 15 groups, 54 trials each (45 second duration), at 55Hz, interpolated via spline to 100Hz, of the movement of their dominant hand, on x axis captured via LEAP sensor. MAT format.

Anonymised Emotional Contagion scores, matched to the participant by unique ID. CSV format.
Anonymised SAM ratings of the stimuli received for each trial, Valence and Arousal ratings. CVS format.
Handedness questionnaire score, Edinburgh Oldfield (1971). CSV format.
PANAS questionnaire scores. CSV format.
The pseudorandomised order of the trial and lags for the acoustic stimuli. Sample of acoustic stimuli used in this experiment. CSV format.
GitHub link for the processing pipelines: https://github.com/EnTimeMent/Spook-and-Play

7 Hindustani raga and singer classification using pose estimation *(Durham)*

Clayton, M., Li, J., Clarke, A. R., Weinzierl, M., Leante, L. & Tarsitani, S. (2021). Hindustani raga and singer classification using pose estimation. OSF. October 14. https://doi.org/10.17605/OSF.IO/T5BWA

7.1 Description

7.1.1 Brief description

The published 'Hindustani raga and singer classification using pose estimation' corpus comprises demonstration performances of nine Hindustani ragas by three professional singers – Apoorva Gokhale, Chiranjeeb Chakraborty and Sudokshina Chatterjee – in the form of solo, unaccompanied alaps; duo recordings of the same nine ragas by Sudokshina Chatterjee with tabla accompaniment; and clips of concert performances by the other two singers (for SCh, we refer to a recording in another OSF corpus). This collection also includes a Colab notebook and python library used to extract and post-process 2D and 3D pose data from the video with the help of third-party pose extraction algorithms (Cao et al. 2021, Mehta et al. 2018); the resulting data, both raw (as JSON) and post-processed (as CSV); and 'output' video clips featuring the movement data or predicted raga or singer classes from the analysis featured in the paper 'Hindustani raga and singer classification using 2D and 3D pose estimation from video recordings' (Clayton, Li, Clarke & Weinzierl, submitted).

Table 6. Summary of contents. The collection comprises the following:

0_original_videos	All solo, duo and concert raga videos, as recorded. A complete list of these clips is saved as Extracts.xlsx.
1_overlay_videos	All solo, duo and concert raga videos, with final movement data (2D skeleton, upper body parts from OpenPose) overlaid
2_blank_3d_video	3D movement data visualised with black background for all solo, duo and concert raga videos
3_json_files	JSON files containing the raw output of 2D and 3D pose estimation algorithms (before post-processing)
4_csv_files	CSV files containing the post-processed 2D and 3D movement data, single file per take per musician
5_output_video	Solo, duo and concert videos with predicted labels for singer and raga printed in the top-left corner
GitHub: DurhamARC/raga-pose-estimation	CoLab notebook and raga pose estimation python library
Documents	Song texts from duo recordings, transcribed by Sudokshina Chatterjee Extracts.xlsx (see above)

5.2 Parameters and initial values used for analysis

The post-processing stage in the generation of movement data involves the selection of two key parameters: (1) the confidence value for each data point below which the value was replaced with the previous value, and (2) the number of frames over which movement is smoothed using the Savitzky–Golay filter. Values for these parameters were established by comparing the output from a range of settings with the original videos and choosing the best match empirically (i.e., the parameters that best captured the actual movement and removed artefacts, but was not smoothed to the point that clearly visible movements were lost). The parameters used in our post-processing are:

Confidence threshold: 0.7

Smoothing window: 13

5.3 Licences

The performers are all professional musicians, and all signed release forms permitting non-profit educational and research use of the recordings. Usage of the recordings is in accordance with the licence attached, which permits use for the purposes of not-for-profit research or teaching or personal educational development.

Usage of the Colab notebook and python library is in accordance with the MIT licence attached to the GitHub repository.

5.4 Ethical approval

Ethical approval. Approval for the recording and sharing of the solo dataset and new concert recordings was granted by the Durham University Music Department Ethics Committee on 17th February 2020 (application MUS-2020-02-03T13_53_06-fghk75). This was extended to cover the duo recordings and their use, approved by the same committee on 19th April 2021 (MUS-2021-03-31T09_31_25-fghk75).

5.5 References

7.1.2 *Publications employing this corpus*

Clayton, M., Li, J., Clarke, A. R., and Weinzierl, M. (submitted). Hindustani raga and singer classification using 2D and 3D pose estimation from video recordings.

7.1.3 *Pose estimation algorithms*

Cao, Z., G. Hidalgo, T. Simon, S. -E. Wei and Y. Sheikh. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (1), pp. 172-186. https://doi.org/10.1109/TPAMI.2019.2929257

Mehta, D., Sotnychenko, O., Mueller, F., Xu, W., Sridhar, S., Pons-Moll, G., & Theobalt, C. (2018). Single-shot multi-person 3d pose estimation from monocular RGB. In *2018 International Conference on 3D Vision*. 120-130. https://arxiv.org/abs/1712.03453v3

8 Qualisys markerless mocap campaign – November 2021

The EnTimeMent Markerless Motion Capture Campaign consisted of a 2-week series of feasibility studies and experiments investigating the recent innovative mocap technology by Qualisys and the novel techniques for movement analysis and prediction developed by the EnTimeMent Consortium. During the two-week campaign data was collected as part of the evaluation studies for the three scenarios of WP 4. Furthermore, other partners of the EnTimeMent project had the opportunity to explore the potential of markerless motion capture for their research projects.

The projects involved in the markerless mocap campaign were:

- Measurement of typical situations for healing and life support for disabled children, UniGe/Gaslini Children hospital (WP 4, scenario 1)

- Capture of pain behaviours (acted by one of the researchers) in a set of everyday type movements at home, with the aim of understanding how well markerless mocap performs in this scenario - UCL (WP 4, scenario 2)
- Leading and following in a group exercise with three persons, Euromov/UniGe (WP 4, scenario 3)
- Measurements of grasping movements for studies of action prediction, IIT
- Measurement of short actions with single and two actors for perception studies (Maastricht University)
- Indian solo singer performance (Durham University)

The recorded data consists of:

- Synchronous video recordings of a multi-camera video system (12 cameras, HD, frame rate: 50-60 Hz). The camera system was calibrated and synchronized.
- Skeleton pose data obtained by analysis with Theia3D markerless software.
- Possible export formats of skeleton pose data:
    - C3D format, readable by C-Motion Visual3D software,
    - MAT format, similar to Qualisys QTM export format readable by Matlab
    - TSV format, text-based format similar to Qualisys TSV export format
    - FBX format, readable by most animation software.

The data is currently being processed. The decision for publication of the data sets will be made per project, depending on the quality of the data and the potential significance of the data set for the community. The progress will be reported in an upcoming deliverable.

For more information about the Markerless Mocap Campaign, see deliverable D3.8.

9 Methodology and dataset description for the project entitled ' Individual limb-dependency in a curvilinear rhythmic movement' (Zarandi, Z., Stucchi, N., Fadiga, L., Pozzo, T.)

7.1 Participants

Forty healthy subjects (14 males and 26 females, two left-handed and 38 right-handed), with normal or corrected-to-normal vision and naive to the task, participated in this study. None of the participants had reported neurological, psychiatric, or other relevant medical problems affecting motor performance. Each participant provided written informed consent before participating in the study. The Ethics Committee of the Regione Emilia Romagna approved the experimental procedures (Comitato Etico di Area Vasta Emilia Centro della Rigione Emilia-Romagna; Ref. EM255-2020 UniFe/170592 EM Estensione).

7.2 Procedure

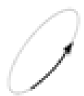
Under an ordinary lighting condition, we organized the whole recordings in a quiet room. The coordinates of drawing movements were recorded via a Wacom tablet (Bamboo slate; temporal resolution: 200 samples/s; resolution: 1748 by 2551; Active area: 210 × 297 mm), placed in a horizontal position in front of the sitting participants.

Subjects were required to perform a simple motor task, drawing ten ellipses continuously on a template ellipse with three different speeds and hands separately. The template ellipse to be traced was drawn on standard-sized (A4) white sheets set on the table. The geometrical features of the template were defined by its eccentricity of 0.968, major semi-axis of 8 cm, and minor semi-axis of 2 cm. In addition, the major axis of the ellipse template was rotated by 45° and -45° for the right and left-hand movements, respectively.

Each participant carried out six various movement conditions considering experimental factors, the drawing speed (3 levels) and the performing hand (2 levels). A metronome indicated the rate of the movements for only spontaneous speed with a period of 80 beats per minute (BPM). For slow and fast speed, subjects produced their preferred slower and faster speed relative to their spontaneous speed. The directions of movement for the right and left hand were counter-clockwise and clockwise, respectively.

Each subject executed ten successive repetitions per condition. In the end, each subject performed a total of 60 repetitions, presented in two blocks, separated by short breaks. The order in which these repetitions were conducted was randomized with the constraint that the same movement type could not occur in the successive trial. At the beginning of the experiment, subjects were asked to synchronize their movement to the beats of a metronome and try to memorize it as the spontaneous speed.

Table 7. Representation of experimental conditions in the hand-drawing task.

Independent variables			Number of trials	Number of repetition in each trial
Hand side (2 levels)		Speed (3 levels)		
Dominant hand: moving in counter-clockwise direction		Slow	10 trials for each condition (totally 60 trials)	10 continuous ellipse in each trial (totally 600 ellipse/reps)
		Spontaneous		
		fast		
Non-dominant hand: moving in clockwise direction		Slow		
		Spontaneous		
		fast		

7.3 Measures

For each repetition, the first and two last ellipses were excluded from data analysis. Then kinematics and geometrical parameters of the seven remained traces were extracted. At the end two main data set obtained: dataset 1 included time-series data (e.g., x, y, tangential velocity, radius of curvature and time), dataset 2 included extracted kinematics and geometrical features for three main categories as listed in Table 8:

1. Kinematic features: In addition to each ellipse's duration and peak velocity values, the shape of the velocity profile calculated for outward and inward strokes was also quantified. Precisely, the timing of the motor plan was evaluated by considering the ratio of acceleration duration (AD) to the total movement duration (MD) along with the two successive directions (outward and inward).

2. Velocity-curvature coupling: The relationship between the geometrical form of the movement and its velocity is well characterized by the empirical relation V(t) = KC(t)1-b where: V is the angular velocity, C is the curvature, at time t; K is a constant, named the velocity gain factor, which depends on the type of movement. To compute the adequacy of a two-third power law, two levels of analysis are undertaken (Viviani 1985). The initial level adopts a straightforward approach of analysing each movement cycle, which its outcome are angular velocity and radius of curvature. In the second step, we calculate b (beta) exponent in equation (V(t) = KC(t)1-b).
3. Geometrical features: The global geometric accuracy of the performances was measured considering several parameters: minor semi-axis of the ellipse, eccentricity (elongation of the ellipse, which is calculated by this equation $e = \sqrt{1-(b/a)^2}$, a is major semi-axis, and b is minor semi-axis). Angle of rotation (the inclination of the major axis of the traces to the horizontal plane), and perimeter of the ellipse were also computed.

Table 8. Three categories of features in dataset 2.

Kinematics features	Velocity-curvature coupling	Geometrical features
Max Velocity	Beta exponent	Eccentricity
Mean Velocity	Correlation coefficient	Minor axis
Movement Duration	K velocity gain factor	Major axis
Max Velocity of each stroke	Jerk cost	Perimeter
Mean Velocity of each stroke		Error index
Time to peak velocity		
Duration of each stroke		

7.4 Extracted features

Table 9 lists the descriptive data expressed as mean and standard deviation (SD) of all analysed parameters.

Table 9. Mean and standard deviation of kinematic and geometrical variables.

	Mean						SD					
	ND-S	ND-N	ND-F	D-S	D-N	D-F	ND-S	ND-N	ND-F	D-S	D-N	D-F
MaxV	14.733	29.235	44.302	15.123	28.965	44.842	5.880	6.156	7.255	6.005	6.699	6.767
Mean V	9.701	19.307	28.896	9.888	19.080	29.316	3.759	4.066	4.678	3.921	4.351	4.410
Duration	3.656	1.663	1.082	3.629	1.678	1.042	2.008	0.574	0.227	1.830	0.714	0.192
Beta	0.331	0.339	0.343	0.332	0.338	0.342	0.013	0.011	0.011	0.014	0.009	0.009
K	4.798	9.381	13.689	4.799	9.210	13.855	1.788	1.955	2.172	1.859	2.052	2.028
Corr-Coef (2/3 PL)	0.988	0.989	0.981	0.985	0.992	0.985	0.018	0.016	0.022	0.023	0.011	0.015
Eccentricity	0.964	0.964	0.968	0.968	0.969	0.972	0.011	0.013	0.014	0.008	0.008	0.008
Major axis	7.723	7.704	7.690	7.755	7.621	7.559	0.748	0.556	0.797	0.897	0.588	0.657
Minor axis	2.023	2.008	1.858	1.932	1.876	1.750	0.287	0.301	0.344	0.283	0.247	0.265
Perimeter	33.318	33.225	32.906	33.264	32.647	32.186	3.100	2.174	3.099	3.791	2.451	2.720

10 UCLIC-Runner dataset (Bi, T., Berthouze, N., Singh, A., & Costanza, E. (2019))

A data collection protocol was developed to investigate the feasibility to build systems that can automatically capture the temporal structures of the experience of running and related runners' emotional and physical states (for more details on the study, see D4.7 and related publication (Bi et al., 2019)). A bespoke wearable voice-based Runner app was developed and used together with insole pressure sensors, arm movement sensors, heart rate and skin conductance sensors (Figure 5). The arm movement acceleration and physiological sensors were gathered through the Empatica bracelet. The insole was a bespoke insole with 8 pressure sensors placed as shown in Figure 5.

Runners of different fitness level were asked to run accordingly to their own running plan while wearing the wearable sensors and voice app. Runners were asked to wear the sensors while answering to voice-based questions (from the voice app) every minute during running. The questions asked them to rate their level of exertion (5 level scale), pain level (5 level scale), emotional state (3 level scale) and desire to continue to run or give up (3 level scale).

11 runners (five males, six females) were recruited for this data recording for a total 14 running sessions. The duration of running sessions averaged 34.7 ± 15.1 minutes. A distribution of the labels can be seen in Figure 6. Insoles, arm movement and physiological data together with the self-reported data will be made available after a second publication we are preparing with be out.

Figure 5: Bespoke system: consisting in pressure sensors, head and arm movement sensors, physiological sensors (from Bi et al., 2019 and D4.7)

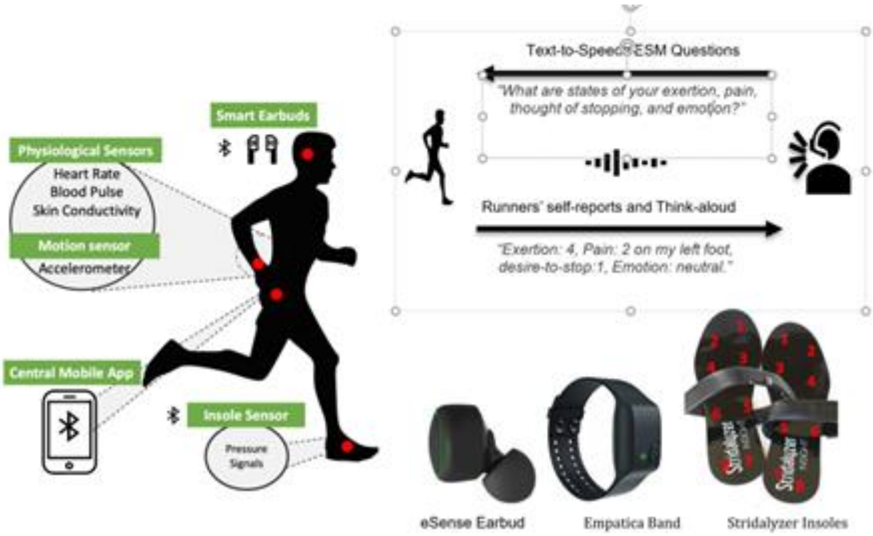
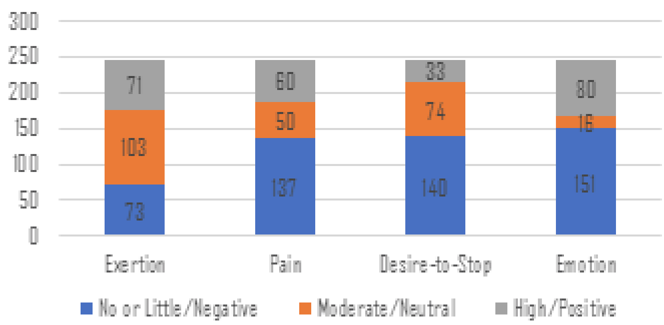


Figure 6. Distribution of self-reported affective and physical state labels collected during running in relation to the sensed data (from D4.7)



11 References

Bi, T., Berthouze, N., Singh, A., & Costanza, E. (2019). Understanding the Shared Experiences of Runners and Spectators in Long-Distance Running Events. *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*. Association for Computing Machinery (ACM). doi:10.1145/3290605.3300691

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Bradley, M. M., & Lang, P. J. (2007). The International Affective Digitized Sounds (; IADS-2): Affective ratings of sounds and instruction manual. *University of Florida, Gainesville, FL, Tech. Rep. B-3*.

Castrillón, M., Déniz O., Guerra C. & Hernández M. (2007). ENCARA2: Real-time detection of multiple faces at different resolutions in video streams. Journal of Visual Communication and Image Representation, 18(2), 130-140. doi: 10.1016/j.jvcir.2006.11.004

Cao, Z., G. Hidalgo, T. Simon, S. -E. Wei and Y. Sheikh. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (1), pp. 172-186. https://doi.org/10.1109/TPAMI.2019.2929257

Clayton, M., Li, J., Clarke, A. R., and Weinzierl, M. (submitted). Hindustani raga and singer classification using 2D and 3D pose estimation from video recordings.

Dalal, N. & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1(1), 886–893. doi:10.1109/CVPR.2005.177

Gaudreau, P., Sanchez, X., & Blondin, J.-P. (2006). Positive and Negative Affective States in a Performance-Related Setting. European Journal of Psychological Assessment, 22(4), 240–249. https://doi.org/10.1027/1015-5759.22.4.240

Issartel, J., Gueugnon, M., & Marin, L. (2017). Understanding the Impact of Expertise in Joint and Solo-Improvisation. Frontiers in Psychology, 8, 1078. https://doi.org/10.3389/fpsyg.2017.01078

Kalal, Z., Mikolajczyk, K. & Matas, J. (2010). Forward-Backward Error: Automatic Detection of Tracking Failures. Proceedings of the 20th International Conference on Pattern Recognition, 2756–2759. doi:10.1109/ICPR.2010.675

Kruppa H., Castrillon-Santana M. & Schiele, B. (2003). Fast and Robust Face Finding via Local Context. Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 157–164. Retrieved from https://accedacris.ulpgc.es/bitstream/10553/17656/5/C022_ICCVW_PETS03.pdf

Lienhart R., Kuranov A. & V. Pisarevsky (2003). Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection. Joint pattern recognition symposium, 2781(1), 297-304. doi:10.1007/978-3-540-45243-0_39

Lucas, B.D. & Kanade, T. (1981) An Iterative Image Registration Technique with an Application to Stereo Vision. Proceedings of the 7th International Joint Conference on Artificial Intelligence, 2(1), 674–679. Retrieved from http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.421.4619

Martin-Soelch, C., Stöcklin, M., Dammann, G., Opwis, K., & Seifritz, E. (2006). Anxiety trait modulates psychophysiological reactions, but not habituation processes related to affective auditory stimuli. *International journal of psychophysiology*, *61*(2), 87-97.

Mehta, D., Sotnychenko, O., Mueller, F., Xu, W., Sridhar, S., Pons-Moll, G., & Theobalt, C. (2018). Single-shot multi-person 3d pose estimation from monocular RGB. In *2018 International Conference on 3D Vision*. 120-130. https://arxiv.org/abs/1712.03453v3

Nardelli, M., Valenza, G., Greco, A., Lanata, A., & Scilingo, E. P. (2015). Recognizing emotions induced by affective sounds through heart rate variability. *IEEE Transactions on Affective Computing*, *6*(4), 385-394.

Plaisant, O., Courtois, R., Réveillère, C., Mendelsohn, G. A., & John, O. P. (2010). Validation par analyse factorielle du Big Five Inventory français (BFI-Fr). Analyse convergente avec le

NEO-PI-R. Annales Médico-Psychologiques, Revue Psychiatrique, 168(2), 97–106. https://doi.org/10.1016/j.amp.2009.09.003

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Olugbade, T. A., Bianchi-Berthouze, N., Marquardt, N., & Williams, A. C. D. C. (2018). Human observer and automatic assessment of movement related self-efficacy in chronic pain: from exercise to functional activity. IEEE Transactions on Affective Computing, 11(2), 214-229.

Olugbade, T. A., et al. (subm.). Human Movement Datasets: An Interdisciplinary Scoping Review.

Ojala, T., Pietikäinen, M. & Mäenpää, T. (2002). Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), 971-987. doi:10.1109/TPAMI.2002.1017623

Osokin, D. (2019). Real-Time 2D Multi-Person Pose Estimation on CPU: Lightweight OpenPose. ICPRAM. Retrieved from https://arxiv.org/abs/1811.12004.

Raymond, A. (2017). La contagion émotionnelle : analyse de l'influence du niveau de stress perçu sur la sensibilité au processus de contagion émotionnelle. Mémoire. Rimouski, Québec, Université du Québec à Rimouski, Unités départementales des sciences de la gestion, 119 p.

Shi, J. & Tomasi, C. (1994). Good Features to Track. IEEE Conference on Computer Vision and Pattern Recognition, 593–600. doi:10.1109/CVPR.1994.323794

Shiqi, Y. (n.d.) Eye Detection. Shiqi Yu's Homepage. Retrieved from http://yushiqi.cn/research/eyedetection.

Słowiński, P., Zhai, C., Alderisio, F., Salesse, R., Gueugnon, M., Marin, L., ... & Tsaneva-Atanasova, K. (2016). Dynamic similarity promotes interpersonal coordination in joint action. *Journal of The Royal Society Interface*, *13*(116), 20151093.

Tomasi, C. & Kanade, T. (1991). Detection and Tracking of Point Features. International journal of computer vision. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.45.5770