The MAMEM Project - A dataset for multimodal human-computer interaction using biosignals and eye tracking information

Spiros Nikolopoulos[1], Kostas Georgiadis[1,2], Fotis Kalaganis[1,2], Georgios Liaros[1], Ioulietta Lazarou[1], Katerina Adam[1], Anastasios Papazoglou - Chalikias[1] , Elisavet Chatzilari[1],Vangelis P. Oikonomou[1], Panagiotis C. Petrantonakis[1], Ioannis Kompatsiaris[1], Chandan Kumar[3], Raphael Menges[3], Steffen Staab[3,4], Daniel Müller[3], Korok Sengupta[3], Sevasti Bostantjopoulou[5], Zoe Katsarou[6], Gabi Zeilig[7], Meir Plotnik[8], Amihai Gottlieb[7,8], Sofia Fountoukidou[9], Jaap Ham[9], Dimitrios Athanasiou[10], Agnes Mariakaki[10], Dario Comanducci[11], Edoardo Sabatini[11], Walter Nistico[12] and Markus Plank[12]

[1]Information Technologies Institute, Centre for Research & Technologies Hellas, Greece, [2]Aristotle University of Thessaloniki, School of Sciences, Department of Informatics, Greece, [3]Institute for Web Science and Technologies, University of Koblenz-Landau, Germany, [4]Web and Internet Science Research Group, University of Southampton, UK, [5]3rd University Department of Neurology, Aristotle University of Thessaloniki, Greece, [6]Department of Neurology, Hippokration General Hospital, Thessaloniki, Greece, [7]Department of Neurological Rehabilitation, Chaim Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel, [8]Center of Advanced Technologies in Rehabilitation, Chaim Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel, [9]Human-Technology Interaction, Eindhoven University of Technology, The Netherlands, [10]Muscular Dystrophy Association, MDA Hellas, Greece, [11]EB Neuro S.p.A, Italy, [12]Senso Motoric Instruments - SMI, Germany

Keywords: BCI, Brain-Computer-Interfaces, Biosignals, eye-tracking, multimodal, human-computer interaction, EEG, GSR, HR, GazeTheWeb, Training, hands-free interaction, disabled, people with disabilities.

Abstract: In this report we present a dataset that combines multimodal biosignals and eye tracking information gathered under a human-computer interaction framework. The dataset was developed in the vein of the MAMEM project that aims to endow people with motor disabilities with the ability to edit and author multimedia content through mental commands and gaze activity. The dataset includes EEG, eye-tracking, and physiological (GSR and Heart rate) signals along with demographic, clinical and behavioral data collected from 36 individuals (18 able-bodied and 18 motor-impaired). Data were collected during the interaction with specifically designed interface for web browsing and multimedia content manipulation and during imaginary movement tasks. Alongside these data we also include evaluation reports both from the subjects and the experimenters as far as the experimental procedure and collected dataset are concerned. We believe that the presented dataset will contribute towards the development and evaluation of modern human-computer interaction systems that would foster the integration of people with severe motor impairments back into society.

1 Introduction

Multimodal interaction refers to human-computer (HC) systems that allow interaction through multiple channels of data input and output apart from the natural modes of communication. One of the most challenging such systems are the Brain-computer interfaces (BCIs) that allow communication without the need of any physical, muscular activity. Instead, they enable the control of a computer using brain signals measured with various recording techniques. A non-invasive and widely used technique is electroencephalography (EEG). Different control components from EEG signals, such as event-related potentials (ERPs) (Handy, 2005) and sensorimotor rhythms (SMR) (Arroyo et al., 1993) have been gaining attention in recent years leading to striking advancements of BCI systems.

However, BCI systems typically rely on supervised classification frameworks of the abovementioned components and, as a result, great volumes of data are needed for reliable development of respective systems. There are numerous publicly available datasets, e.g., various datasets have been released in the context of the BNCI H2020 project[1]. Nevertheless, the majority of the available datasets incorporate only EEG signals and very few bimodal signals such as EEG and EOG (electrooculography) or ECoG (electrocorticography). Very few datasets contain multimodal signals in addition to brain activity such as physiological signals, e.g., electro-dermal activity (EDA, frequently measured as Galvanic Skin Response (GSR)) and heart rate (HR), or eye tracking information. Additional modalities could further enhance the reliability of an HCI system, e.g., using eye-tracking information for more complex controls of the computer, and simultaneous multifaceted evaluation of the interaction, e.g., assessing the stress levels of the user's experience of the HCI system usage via physiological signals. Besides, there is lack of multi-subject and multi-modal datasets that are acquired during specific applications and usage scenarios, such as multimedia editing, web browsing or texting in social media. Such datasets would further demonstrate the applicability and robustness of the EEG components and foster the development of more practical BCI systems.

The aim of this report is to describe the process of collection and archiving of a large, multimodal and multi-subject dataset gathered in the context of the MAMEM project[2]. MAMEM's goal is to integrate people with motor disabilities back into society by endowing them with the critical skill of managing and authoring multimedia content using novel and more natural interface channels. These channels are controlled by eye-movements and mental commands/brain? (expressed through EEG components), significantly increasing the potential for communication and exchange in leisure (e.g., social networks) and non-leisure context (e.g., workplace). In this vein, a large dataset of EEG, GSR, HR, and eye-tracking signals have been recorded under different experimental protocols where 34 subjects (18 healthy and 16 patients) are using a novel, eye-tracking based, web browsing system, namely the GazeTheWeb browser (Kumar et al., 2017). The dataset was created by gathering signals both from healthy subjects and subjects with motor impairments. The dataset is intended to be exploited by researchers working on HCI and BCI systems' development and more particularly on multimedia editing and authoring related applications.

The paper is structured as follows: in section 2 the demographics of the subject cohorts are presented. Section 3 refers to the materials and methods adopted during the experiments whilst section 4 elaborates on the experimental protocols. In section 5, the signal preprocessing and archiving procedures are described in detail whereas section 6 discusses evaluation issues regarding the experimental procedures

[1] http://bnci-horizon-2020.eu/
[2] http://www.mamem.eu/

and strategies followed. Finally, section 7 concludes the paper.

2 Subjects

2.1 Cohorts and Settings

The sample comprised a group of 36 individuals (N = 36). There were 18 participants with motor impairment, namely, 6 with Parkinson's disease (PD), 6 with Neuromuscular disorders (NMD) and 6 with Spinal Cord Injuries (SCI) (of which 2 dropped out). Moreover, there were 18 able-bodied participants matched in socio demographic profile of patients in all clinical sites[3]. The rationale of recruiting able-bodied participants was to assess MAMEM's feasibility and usability among able–bodied individuals free from any medical condition that may affect brain-computer interfaces and gaze behavior analysis. Then, 18 patients (6 patients of each clinical condition), age- and gender- matched to the healthy participants were recruited. Participants with PD were over 55 (mean 64 years old), while NMD and SCI participants were much younger, since in both conditions the prevalence is in young ages. Demographics and clinical characteristics of the participants appear in the **Tables 1, 2**. In total 27 male and 9 female subjects took part in the study. None of both subjects and controls had previously participated in a BCI study. The study was carried out in accordance with the Declaration of Helsinki and was approved by the institutional ethical committees from each clinical site (AUTH 325, 13/07/2016 MDA HELLAS TH.COM-23, 26/01/2017 and SHEBA v4 3268-16-SMC, 27/12/16). Informed consent was obtained from all the participants after receiving a full explanation of the goals and protocol of the study.

Settings: The clinical trial took place from the beginning of February 2017 to 30 of March 2017 in three different clinical settings. More specifically, PD participants were recruited from the 3rd Neurological Clinic of the Aristotle University of Thessaloniki (AUTH), Greece[4], NMD participants from the MDA Hellas[5] (MDA) and SCI from the rehabilitation center, SHEBA Medical Center[6].

Table 1 Characteristics of the participants

Participant ID	Gender	Age	Education (years)	Diagnosis	time after disease/trauma (years)	Wheel chair user (yes/no)
PP1 (AUTH)	m	71	17	PD	2	0
PP2 (AUTH)	m	56	17	PD	10	0
PP3 (AUTH)	m	67	15	PD	16	0
PP4 (AUTH)	m	71	18	PD	7	0
PP5 (AUTH)	m	57	9	PD	10	0
PP6 (AUTH)	f	62	16	PD	16	0
PH1 (AUTH)	m	48	12	Able-bodied	-	0

[3] There were three clinical sites, Aristotle University of Thessaloniki (AUTH), MDA Hellas (MDA), and Sheba Medical Centre (SHEBA)
[4] http://www.med.auth.gr/
[5] http://www.mdahellas.gr/
[6] https://eng.sheba.co.il/

PH2 (AUTH)	m	56	16	Able-bodied	-	0
PH3 (AUTH)	m	73	18	Able-bodied	-	0
PH4 (AUTH)	f	46	12	Able-bodied	-	0
PH5 (AUTH)	m	52	18	Able-bodied	-	0
PH6 (AUTH)	m	45	18	Able-bodied	-	0
NP1 (MDA)	m	35	16	SMA III	34	1
NP2 (MDA)	m	44	6	Muscular Dystrophy	39	1
NP3 (MDA)	m	32	24	Muscular Dystrophy type II	31	1
NP4 (MDA)	f	36	16	Tunesian Muscular Dystrophy	33	1
NP5 (MDA)	m	25	16	Duchene Muscular Dystrophy	23	1
NP6 (MDA)	f	33	16	Tunesian Muscular Dystrophy	20	1
NH1 (MDA)	f	46	12	Able-bodied	-	0
NH2 (MDA)	f	31	16	Able-bodied	-	0
NH3 (MDA)	m	40	16	Able-bodied	-	0
NH4 (MDA)	m	43	12	Able-bodied	-	0
NH5 (MDA)	f	39	15	Able-bodied	-	0
NH6 (MDA)	m	29	16	Able-bodied	-	0
SP1 (SHEBA)	m	37	23	C3 complete	14	1
SP2 (SHEBA)	m	25	12	C5 incomplete	1	1
SP3 (SHEBA)	m	62	16	C3 incomplete	40	1
SP4 (SHEBA)	m	51	12	C4 incomplete	30	1
SP5 (SHEBA)	m	64	22	C5 complete	20	1
SP6 (SHEBA)	f	31	18	C5 incomplete	13	1

SH1 (SHEBA)	m	30	16	Able-bodied	-	0
SH2 (SHEBA)	m	51	16	Able-bodied	-	0
SH3 (SHEBA)	m	60	13	Able-bodied	-	0
SH4 (SHEBA)	m	63	25	Able-bodied	-	0
SH5 (SHEBA)	f	25	15	Able-bodied	-	0
SH6 (SHEBA)	m	32	17	Able-bodied	-	0

* PPx: Parkinson cohort - Patient, PHx: Parkinson Cohort - Healhty

** NPx: Neuro-muscular disorders cohort - Patient, NHx: Neuro-muscuular disorders cohort - Healthy

*** SPx: Spinal-cord injury cohort - Patient, SHx: Spinal-cord injury cohort - Healthy

2.2 Inclusion/exclusion criteria

Various inclusion and exclusion criteria were utilized for the final selection of the cohorts that participated in our experiments. At any case it was at the participant's discretion as to whether she/he wishes to participate in the investigation activity or not. Neurologists from three different clinical sites, who specialize in motor impairments recruited three different cohorts of patients, interviewed them for initial screening of eligibility for inclusion and provided them additional information about the trial. Both women and men were eligible to take part in the study. It was of primary importance to exclude patients who had insufficient understanding of the study's goal. In order to carry out this trial, we finally selected 36 subjects, following specific inclusion and exclusion criteria:

Inclusion Criteria:

a) *for SCI participants*: neurological level from C5 and below, AIS A,B,C according to the International Standards for Neurological Classification of Spinal Cord Injury (ISNCSCI) examination developed by the American Spinal Injury Association (Kirshblum et al., 2011).

b) *for PD participants*: i) between 45-75 years old and ii) patients were diagnosed with Parkinson's disease according to the UK Parkinson's Disease Society Brain Bank Diagnostic Criteria (Hughes et al., 1992), iii) have bradykinesia and rest tremor of the upper extremities.

c) *for NMD participants*: i) between 18-35 years old and have been diagnosed with one type of NMD (Duchenne, SMA, SMA II, Tunesian, ALS, Arthrogryposis), ii) do not have cognitive impairments.

All patients fulfilled the basic inclusion criteria for the study, i.e., permanent motor impairment or progressive motor impairment with future inability to operate computer and had the ability to understand the instructions.

Exclusion Criteria:

a) *for PD participants:* i) atypical Parkinsonism and ii) those who had received medication for other neurological disorders or had predominant disabling comorbidity, iii) those with presence of involuntary eye movements, iv) any psychiatric or cognitive condition that may interfere with understanding the instructions (including PD patients with dementia), v) implanted devices (DBS and other electrical medical devices) that may interfere with the brain electrical activity recorded by the EEG sensor, vi) brain disorders, such as trauma, stroke, surgery etc., that may interfere with the brain electrical activity

recorded by the EEG sensor, vii) diminished visual acuity, viii) very severe body involuntary movements/dyskinesias for PD patients, ix) medical conditions that may induce seizures, prominent EEG abnormalities, such as continuous slowing, epileptiform discharges.

b) *for NMD and SCI participants:* i) Involuntary eye movements, ii) implanted devices or brain conditions such as brain trauma, brain surgery, stroke that may interfere with the brain electrical activity recorded by the EEG sensor, iii) Medical conditions that may induce seizures, iv) Any psychiatric (e.g., major depression) or cognitive conditions that may interfere with understanding the researchers' explanations, v) Bedridden and vi) Drugs or alcohol abuse.

Withdrawal criteria:

Participants were informed that they are free to withdraw from participation at any point, that their personal data will remain confidential and anonymous, that collected data will be analysed both individually and for the entire group of participants and that securing their privacy and anonymity is of high priority. In case the participants couldn't understand the instructions of the study's personnel, were unable to complete at least 50% of the protocol or could not use the MAMEM platform (e.g., unable to control the computer with brain or eyes after the 1-2 hours of practice) or if there was lack of cooperation with the study's personnel, researchers had the authority to exclude participants from the study.

2.3 Demographics

As previously indicated, in our study there were 18 able-bodied participants [13 male, Age: M(SD)= 45 (13), Education: M(SD)= 15.7 (3.14)] in the sample matched with the socio-demographic profile of 18 motor-impaired participants [14 male, Age: M(SD)= 47.7 (9.75), Education: M(SD)= 16.04 (4.58)]. In **Table 2**, means and standard deviation of demographic characteristics for each group of motor-impaired and their matched able-bodied participants are listed, as well as the ratio of female and male participants. Also, detailed description is provided regarding marital status, number of children, type of working and hand preference per cohort.

Table 2 Distribution of socio-demographic characteristics of participants by groups and in total

	Able - bodied		**PD**		**Total**		**Able - bodied**		**NMD**		**Total**		**Able-bodied**		**SCI**		**Total**	
	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)	N	**% /** mean (std)
Age	6	53.33 (10.5)	6	64.0 (6.7)	12	58.7 (10.1)	6	38.0 (6.7)	6	34.2 (6.2)	12	36.1 (6.5)	6	43.5 (16.5)	6	45 (16.4)	12	44.2 (15.7)
Education	6	15.7 (2.9)	6	15.3 (3.3)	12	15.5 (3.0)	6	14.5 (2.0)	6	15.7 (5.7)	12	15.1 (4.1)	6	17 (4.1)	6	17.1 (4.7)	12	17 (4.2)
Gender																		
Male	5	83.8	5	83.8	10	83.8	3	50	4	66.7	7	58.3	5	83.8	5	83.8	10	83.8
Female	1	16.7	1	16.7	2	16.7	3	50	2	33.3	5	41.7	1	16.7	1	16.7	2	16.7
Marital Status																		
Single	0	0	0	0	0	0	3	50	5	83.3	8	66.7	3	50	2	33.3	5	58.3
Married	6	100	6	100	12	100	2	33.3	1	16.7	3	25	3	50	4	66.7	7	41.7
Divorced	0	0	0	0	0	0	1	16.7	0	0	1	8.3	0	0	0	0	0	0
Children Number																		
0	1	16.7	0	0	1	8.3	4	66.7	6	100	10	83.3	3	50	3	50	6	50
1	0	0	3	50	3	25	1	16.7	0	0	1	8.3	0	0	2	33.3	2	16.7
2	5	83.3	3	50	8	66.7	1	16.7	0	0	1	8.3	2	33.3	0	0	2	16.7

3	0	0	0	0	0	0	0	0	0	0	0	0	1	16.7	1	16.7	2	16.7
Working																		
Full time	6	100	2	33.3	8	66.7	5	83.3	3	50	8	66.7	4	66.7	5	83.3	9	75
No	0	0	4	66.7	4	33.3	1	16.7	3	50	4	33.3	2	33.3	1	16.7	3	25
Hand preference																		
Right	6	100	6	100	12	100	6	100	6	100	12	100	5	83.3	5	83.3	10	83.3
Left	0	0	0	0	0	0	0	0	0	0	0	0	1	16.7	1	16.7	2	16.7

2.4 Clinical characteristics

All patients were neurologically tested for their clinical condition and impairment status by special neurologists. More specifically, they were examined for bradykinesia and immobilization in specific parts of their body and face. The level of motor impairment was rated as follows: a) 0: in case of absence of any symptom, b) 1: in case of partial impairment and c) 2: in case of complete motor impairment in the respective part. Distribution of clinical characteristics of PD, SCI and NMD patients are shown in **Table 3.**

Table 3 Distribution of clinical characteristics of participants –Bradykinesia and Immobilization

	Tongue		**Jaw**		**Neck**		**Shoulders**		**Arms**		**Elbows**		**Wrists**		**Hands**		**Fingers**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
PD																		
Bradykinesia																		
No Symptom	0	0.0	0	0.0	2	33.3	1	16.7	1	16.7	0	0.0	0	0.0	0	0.0	0	0.0
Partial	0	0.0	0	0.0	4	66.7	5	83.3	4	66.7	2	33.3	1	16.7	1	16.7	1	16.7
Complete	0	0.0	0	0.0	0	0.0	0	0	1	16.7	4	66.7	5	83.3	5	83.3	5	83.3
NMD																		
Immobilization																		
No Symptom	4	66.7	4	66.7	2	33.3	1	16.7	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Partial	2	33.3	2	33.3	3	50.0	3	50.0	4	66.7	3	50.0	3	50.0	3	50.0	3	50.0
Complete	0	0.0	0	0.0	1	16.7	2	33.3	2	33.3	3	50.0	3	50.0	3	50.0	3	50.0
SCI																		
Immobilization																		
No Symptom	0	0.0	0	0.0	4	66.7	3	50.0	1	16.7	1	16.7	1	16.7	0	0.0	0	0.0
Partial	0	0.0	0	0.0	2	33.3	1	16.7	2	33.3	2	33.3	2	33.3	3	50.0	3	50.0
Complete	0	0.0	0	0.0	0	0.0	2	33.3	3	50.0	3	50.0	3	50.0	3	50.0	3	50.0

1. **Parkinson's Disease (PD):** Six patients with Parkinson's disease (PD) and six able - bodied controls participated in the clinical trial. Patients with Parkinson's disease had disease duration of 10.2±5.4 years,

age at disease onset of 58.83±10.1 years and they were on 2.2±0.4 stage of disease (Hoehn and Yahr stage). Clinical information and remarks regarding tremor and dyskinesia of PD cohort are provided in **Table 4** and **Table 5** respectively**.**

Table 4 Distribution of clinical characteristics of PD participants –Tremor

	Tongue		Jaw		Neck		Shoulders		Arms		Elbows		Wrists		Hands		Fingers	
Tremor	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
No Symptom	6	100.0	5	83.3	6	100.0	5	83.3	2	33.3	1	16.7	0	0.0	0	0.0	0	0.0
Mild/Mod	0	0.0	1	16.7	0	0.0	1	16.7	3	50.0	4	66.7	4	66.7	4	66.7	5	83.3
Severe	0	0.0	0	0.0	0	0.0	0	0.0	1	16.7	1	16.7	2	33.3	2	33.3	1	16.7

Table 5 Distribution of clinical characteristics of PD patients- Dyskinesia

	Tongue		Jaw		Neck		Shoulders		Arms		Elbows		Wrists		Hands		Fingers	
Dyskinesia	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
No Symptom	0	0.0	0	0.0	5	83.3	5	83.3	4	66.7	4	66.7	4	66.7	4	66.7	4	66.7
Mild/Mod	0	0.0	0	0.0	1	16.7	1	16.7	2	33.3	2	33.3	2	33.3	2	33.3	2	33.3
Severe	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

2. Neuromuscular Disorders (NMD): In our study six patients with a diagnosis of one type NMD took part. More specifically, one with SMA III, two with Tunesian Muscular Dystrophy, two with Muscular Dystrophy and one with Duchene Muscular Dystrophy. All of them use wheelchair and required constant support by nursing staff. All wheelchairs except two were electrically powered. All patients had poor residual muscular strength either of proximal or distal arm muscles. Also, all patients required a mechanical support to maintain neck posture. Finally, all patients retained effective eye movement control. Prior to the study, no patient used technologically advanced aids **(Table 6).**

Table 6 Distribution of clinical characteristics of NMD participants

	N	%
Wheel chair - type		
motorized	4	66.7
regular	2	33.3
Car		
Yes	4	66.7
No	2	33.3
Drive		
No	6	100.0
Hours in bed		
5	1	16.7
8	1	16.7
10	2	33.3
11	1	16.7
12	1	16.7
Financial support		
Financial support from state	5	83.3
Salary/Paralympic	1	16.7
Diagnosis		
Duchene Muscular Dystrophy	1	16.7
Muscular dystrophy	1	16.7
Muscular dystrophy type II	1	16.7
SMA III	1	16.7
Tunesian Muscular Dystrophy	2	33.3
Self movement		
No	4	66.7
Yes	2	33.3

Rehabilitation		
No	6	100.0

3. Spinal Cord Injury (SCI): The SCI participants were six adults, one woman and five men, aged 25–64. Two of the men had spinal cord injuries (three at C5, two at C3 and one at C4) and were confined to wheelchairs, five with motorized and one with regular (**Table 7**).

Table 7 Distribution of clinical characteristics of SCI participants

	N	% / mean (standard deviation)
SCI severity		
C3 complete	1	16.7
C3 incomplete	1	16.7
C4 incomplete	1	16.7
C5 complete	1	16.7
C5 incomplete	2	33.3
SCI cause		
Assault	1	16.7
Fall	1	16.7
Transport	4	66.7
Wheel chair - type		
Powered	5	83.7
Manual	1	16.7
Own a Car		
Yes	5	83.7
No	1	16.7
Drive		
No	5	83.7
Yes	1	16.7
Hours in bed (per day)	6	10.3 (4.6)
Financial support		
VA	3	50
Work	1	16.7
Social security	2	33.3
Months in Rehabilitation	6	12.3 (5.8)

2.5 Medication

Medication treatment of all participants was made by neurologists/movement disorders specialists.

Parkinson Disease: In our study all patients were treated with levodopa and dopaminergic agonists (3 with pramipexole, 2 with piribedil and 1 with rotigotine) on a daily basis. Furthermore, 3 patients were on entacapone,1 on amantadine,1 on biperiden and 2 on antidepressant drugs (mirtazapine and escitalopram). Medication treatment of each participant can be found in **Table 8**.

Neuro-Muscular-Disorder: Two patients were treated with propranolol that is a medication of the beta blocker type and it is used to treat high blood pressure and a number of types of irregular heart rate. One patient was treated with irbesartan which is an angiotensin II receptor antagonist used mainly for the treatment of hypertension. Also, one patient was taking on a daily basis carvedilol for treating high blood pressure, losartan, which is also an angiotensin II receptor antagonist drug used mainly to treat high blood pressure (hypertension) and deflazacort, a glucocorticoid used as an anti-inflammatory and

immunosuppressant. Two out of six participants were taking no medication (**Table 8**).

Spinal Cord Injury: Three out of six participants were taking tolterodine, an antimuscarinic (anticholinergic) agent, for controlling bladder or treating symptoms of urinary frequency, urgency, and leakage. The same three patients were also taking senna on a daily basis for bowel management. Two out of six SCI participants were treated with diazepam and escitalopram to ameliorate anxiety and depressive symptoms respectively. Moreover, two patients were taking magnesium and two calcium, for keeping blood pressure normal, bones strong and the heart rhythm steady. Also, three out of six were taking vitamin D3, which plays a key role in muscle function and the immune system. One of our SCI patients was taking Dantrolene sodium, a postsynaptic muscle relaxant. Furthermore, another SCI patient was treated with aspirin, atorvastatin and a combination of Vitamin C and D, while two others were taking omega 3, nutrients which are helpful and important for circulatory system and offer protection on the heart. Finally two out six were treated with Vitamin B12, which boosts mental function and the immune system (**Table 8**).

Table 8 Medication Treatment of participants with Parkinson Disease, Neuromuscular Disorders and Spinal Cord Injury

ID	Medication	ID	Medication	ID	Medication
PP1	Levodopa, Pramipexole	NP1	-	SP1	Baclofen, Pregabalin, Magnesium, Warfarin, Senna, Vit.C, Nitrofurantoin, Tolterodine Tartrate, Zoledronic acid
PP2	Levodopa, Eescitalopram, biperiden, Piribedil	NP2	Irbesartan	SP2	Senna, Tolterodine Tartrate, Magnesium, calcium, Vit B12, Vit. D3
PP3	Levodopa, Entacapone, Pramipexole	NP3	-	SP3	Diazepam, Dantrolene Sodium, Vit D3
PP4	Levodopa, Eentacapone, Rotigotine	NP4	Propranolol	SP4	Tolterodine Tartrate, Senna, Calcium, Vit B12, Vit. D3,
PP5	Levodopa, Piribedil, Mirtazapine	NP5	Carvedilol, Losartan, Deflazacort	SP5	Aspirin, Atorvastatin, Vit C+D, Omega 3, Q10
PP6	Levodopa, Entacapone, Piribedil, Amantadine	NP6	Propranolol	SP6	Omega3, Vit. C, Vit. D3, Escitalopram

2.6 Consent form and ethical approvals

Subjects were informed regarding the general features and aims of the study, which was approved by the ethics committee of each clinical site and met the standards of the Declaration of Helsinki[7]. Subjects were given a deep explanation and asked some questions about the trial to ascertain that they understand all issues involved in this specific research. In particular, an interactive discussion with the patients and their relatives allowed researchers to provide detailed information and descriptions/specifications of: a) purpose of the research, b) duration of the research activities, c) adopted procedures, d) voluntary participation, e) possible risks, discomfort or disadvantages, f) benefits to the subject or others, d) data protection and confidentiality, and privacy policies, and h) what happens to data, samples and results at

[7] WMA Press Release: WMA revises the Declaration of Helsinki. 9 October 2000

the end of the research.The informed consent process is central to the ethical conduct of research and all participants (and their relatives when required) gave their written informed consent. Moreover, a policy of strict compliance with the trial protocol was adopted.

3 Materials and methods

3.1 Rationale of the trials

The overall objective of the clinical trials, from which the data presented in this report were obtained, was to assess the usability of the developed assistive technology by evaluating the different assumptions and options for its implementation. In designing and deploying any new technology, training plays a central role in influencing user acceptance as it is the first interaction between a user and the technology. Furthermore, user training is required so that the technology assessment results can be reliably attributed to the pros and cons of the technology instead of the user's familiarity with the system.

In this regard, the clinical trials have been designed to include a training and a testing cycle. In the former, users are introduced to the platform, the application of the EEG headset and the eye tracker, and provided with the basic skills to operate the MAMEM system. The latter refers to multimedia managing, authoring and sharing using dictated tasks such as writing an e-mail and photo editing.

In designing an efficient training procedure, the training cycle includes introductory exercises scoping to present to the user the basic components of controlling a computer with different sensors. For this, we break down the general task of controlling a computer with eyes and mind in different components and let the learner practice the individual components. In order to maximize learning effectiveness, the training starts with easy tasks and builds them slowly up to more difficult tasks. The training is therefore divided into three main parts: **a) Basic:** includes the basic eye-tracking functionality by training the user on focusing on various locations. **b) Intermediate:** includes the training of the user with respect to gaze-based interface and its buttons, as well as the basic functionalities of EEG (ErrPs and SMR system calibration and user training). **c) Advanced:** includes the introduction of the user to the advanced functionalities of the system, such as the system settings, word prediction, etc. Each task has different sublevels, again building up from easy to difficult.

Next, to allow the user to have enough skills to make use of the MAMEM system, a mastery learning system was used. In this mastery learning system, before users can proceed to the next level, they have to possess a specified mastery of content knowledge and skills. This is implemented in two different ways. First, users are advised to redo a level if their score is on the low side. Secondly, users acquire new skills in the training cycle, and apply these skills during the testing cycle (with the dictated tasks), where they will have to apply the skills for the operation of programs. Four dictated tasks were performed during the testing cycle of the trials: a) Writing and sending an email, b) Editing a photo online, c) Making a tweet on Twitter, d) Finding and watching a video on YouTube.

The dataset presented in this work was gathered during the procedures described above concerning different experimental cycles and difficulty levels.

3.2 Persuasive vs Non-persuasive & Gamification

Persuasive technologies (Fountoukidou et al., 2016; Fountoukidou et al., 2017) are used to convince, stimulate and motivate users to engage in various behaviours. A critical goal was to use persuasion design strategies in the training process of the MAMEM system, to present the most important insights into how to motivate individuals with disabilities in learning and operating this system and to use it for increasing

their social participation. This was realized as a plethora of tailored strategies, taking into consideration the needs and requirements of each of the three groups of MAMEM users (SCI, PD, NMD). Two versions of the training software were created to facilitate the experimental process. The first one included persuasive design strategies, and the second one was stripped to the bare minimum game elements with an absence of persuasion strategies. The persuasive design strategies had two different objectives: i) user acceptance and engagement, and ii) social inclusion. At the end of the experiment, helpful insights were extracted as to how a persuasive design related system is better than a stripped one at motivating users to use a new platform or enhancing the rate in which the user learns something new.

Half of both able bodied and disabled participants used the training software (see section 3.3) with the inclusion of the persuasive design and the other half of them the stripped one. This provided data with regard to their performance and their learning curve (i.e., time and errors) and allowed a comparison and indication of the effectiveness of the persuasive design elements. Learning curves are useful learning performance measures, in extracting information on when users exhaust their capacity to learn. In general, learning curves have an initial slow beginning, followed by steep acceleration and an ultimate plateau of optimal performance. In the final phase, execution time and error rate decrease linearly with the skill level of the user. In case execution time and/or error rate do not further decrease, users can be assumed to have reached a 'plateau' in their capacity to improve their performance.

In both versions, the process of training the end-users on operating the system has been gamified, breaking it down into levels of varying difficulty (see section 3.1) so as to accommodate both for users without prior experience and users who have used assistive technologies in the past. Users could only proceed to the next level after they had acquired a certain skill, which was ensured by prompting them to redo a level until they reached a sufficient score, based on completion time and failures.

Further gamification elements were integrated in the training experience of the first version of the training software, including trophies, a way to recognize player accomplishments within a social group, scoreboards, comparing performance against previous runs, leaderboards, comparing performance between players and assignments, short term objectives shaping the gameplay narrative. The implementation of persuasion strategies resulted in tailoring the training experience to each user, by offering: a) personalization, addressing the player using its username, as well as using demographics to create personalized content, like ethnicity specific quiz questions, b) descriptive feedback, providing prompt messages depending on the progress from previous game levels and c) content matching, providing user-specific evaluative feedback, making the users feel that the game is tailored to their personality.

Also the context of evaluative feedback that the training web application provided to the users was varying based on the subject group type, like for example, male or female and young or older users. In young male and female users, the evaluative feedback of the performed tasks was in form of praise for completion time and scores, or acknowledgement of accomplishments. For older male users the feedback provided was more encouraging in the case of failing a task, and for older female ones the feedback provided highlighted social aspects like the perception of family members or the patient’s doctor in case of a task’s success or failure.

3.3 GazeTheWeb

MAMEM platform is centred on interactive web access for end-users by novel input mechanism. In this regard, GazeTheWeb works as the primary framework to support web interaction by non-conventional input sources to enhance web accessibility. In the trials, the above-mentioned dictated tasks were

performed using GazeTheWeb.

GazeTheWeb (Menges et al. 2017) has been developed as part of MAMEM project and uses gaze signals obtained from eye tracking devices to control the web application environment, and supports full-featured web browsing with eye-based input mechanism. It integrates the visual appearance and control functionality of webpages in an eye tracking environment, i.e., the input events which are typically composed of mouse and keyboard interactions in generic applications. In GazeTheWeb, the problem of gaze-based browsing has been resolved by a unifying framework of web engineering and interface design, i.e., identification of the interactive elements of webpages (e.g., text input, hyperlinks, scrollable sections, edit box, etc.) via unsupervised extraction of Document Object Model (DOM) nodes, and representation of these elements with explicit/implicit indicators to be accessed by eye gaze inputs. The Web browser environment is built upon the Chromium Embedded Framework[8] (CEF), which provides an appropriate architecture to enhance the web with interfaces for gaze interaction (Kumar et al. 2016).

Figure 1 shows the interface design of GazeTheWeb browser, which can be divided into three components. On the left, the Web panel covers common actions like going back and forward or opening the tab overview of the browser. The Tab panel on the right side of the interface represents the possible actions on the current webpage, like text selection, click emulation or to enable automatic scrolling. The central part of the interface is the Webview containing the actual webpage, rendered by the underlying Chromium project. Interactive overlays are added for input fields within the page. The T-letter inside a circle representing the text input field to submit query on DuckDuckGo search page.

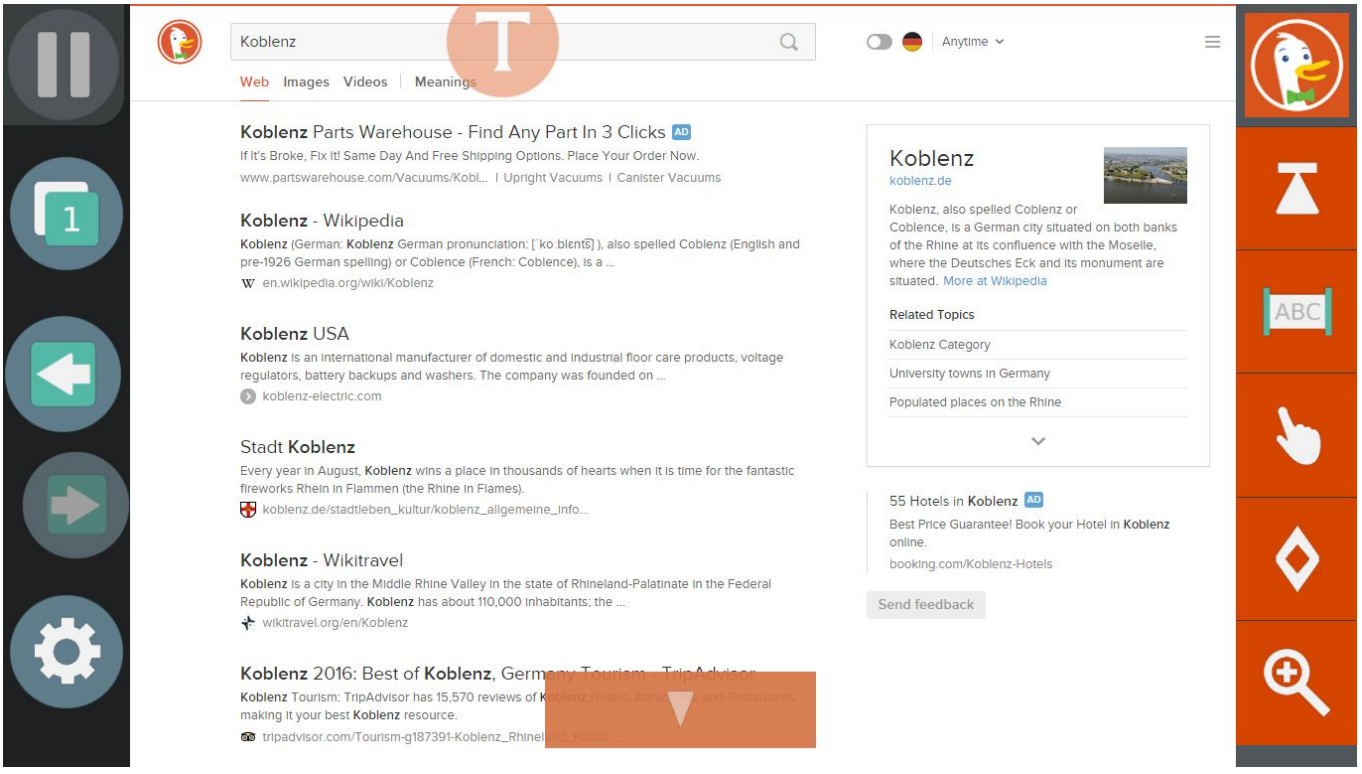

Figure 1: GazeTheWeb Interface - Web Panel, Web View and Tab Panel

GazeTheWeb supports the efficient interaction with the webpage (scrolling, link navigation, text input), and necessary browsing menu operations like bookmarks, history, and tab management, which is essential to perform every-day browsing activities like sending email, photo editing, social media, and video interaction (used in the trials). The detailed functionality of GazeTheWeb can be seen in the online

[8] https://bitbucket.org/chromiumembedded/cef

demonstration[9]. Furthermore the framework is available as an open source system[10], providing in-depth information of the architecture (Kumar et al., 2017).

3.4 Apparatus - Hardware & software

The signals that were captured from the participants during the experimental trials are a) multi-channel EEG data, b) eye-gaze position on the screen based on an eye-tracker device, c) galvanic skin response (GSR) measured by two sensors placed on their index and middle fingers, and d) heart rate in beats per minute (bpm) measured by an optical sensor placed on the tip of their index finger.

The signal capturing platform is divided into two different configurations, a "Heavyweight" configuration consisting of high-end devices able to capture high quality signals and a "Lightweight" configuration that is relatively less reliable than the previous but consists of more affordable consumer-based devices. This separation applies only for the EEG recording and eye-tracker devices since the GSR & OHR sensor is relevant to both configurations, being both an affordable device and a high quality sensor.

The data streams from all four modalities were captured and synchronized with the help of their official SDKs and the LabStreamingLayer (LSL) library[11]. Two different laptops were used for the execution of the experiments. The first one was an Alienware 15 R2 with a i7-6700q CPU and 16GB RAM which covered the experiments that took place at the AUTH and MDA sites. For the SHEBA site, we used a HP Elitebook 840 G1 with a i5 4310U CPU and 8GB of RAM. The complete list of the software that was involved in our experiments have been made available through Github via an installer package[12].

3.4.1 Eye-tracking devices and SDKs

The eye-tracking devices that were selected for use in our study were the **SMI REDn Scientific**[13] eye tracker for the Heavyweight configuration and the **Visual Interaction myGaze n**[14] eye tracker for the Lightweight configuration. The main difference between those devices is the maximum supported sampling rate for capturing the eye-gaze, being 60Hz for the Heavyweight configuration and 30Hz for the Lightweight version.

The two devices provide similar APIs in their SDKs for receiving the raw eye-tracker data. Two simple applications were developed for receiving the data and push them to the LSL. Specifically, two streams consisting of the x & y gaze coordinates in screen pixels were pushed to LSL and recorded in the dataset.

3.4.2 EEG devices and SDKs

For the EEG part, the BePlusLTM Bioelectric Signal Amplifier[15] was used for the Heavyweight configuration, consisting of 61 EEG electrodes, 1 reference and 1 ground electrode (see Appendix 9.1). The reference electrode is placed in the area between FCZ and FZ and the ground between CPZ and PZ. Skin impedance was reduced using a special electro conductive cream and was kept below 10 KΩ for all electrodes.

The Emotiv EPOC[16] device was selected for Lightweight configuration, which is equipped with 14 saline-based, wet-contact resistive electrodes, covering the 10-20 electrode area, fixed to flexible plastic arms of a wireless headset. EPOC is an affordable EEG device targeting the consumer market which has

9 https://youtu.be/x1ESgaoQR9Y

10 https://github.com/MAMEM/GazeTheWeb
11 https://github.com/sccn/labstreaminglayer
12 https://github.com/MAMEM/mamem-platform/releases
13 https://www.smivision.com/eye-tracking/product/redn-scientific-eye-tracker/
14 http://www.mygaze.com/products/mygaze-n/
15 http://www.ebneuro.biz/en/neurology/ebneuro/galileo-suite/be-plus-ltm
16 https://www.emotiv.com/epoc/

also been used by several research groups in BCI applications.

3.4.3 GSR devices and SDKs

The Shimmer3 GSR+[17] device was used in both light and heavy-weight configurations for capturing the GSR and HR signals. Shimmer is a small-sized device that was tied to the participant's hand by a wrist-band. Due to its small size and weight, it was not expected to disturb the participant during the experiments. Shimmer has the capability of capturing many type of signals, such as accelerometer data, Electrocardiogram (ECG), etc. In our case we configured the Shimmer to only capture the GSR and HR of the participant. GSR was measured by placing two electrodes on the primal phalanges of the index and middle fingers. A small voltage is provided on those two electrodes which was then converted in skin resistance levels measured in KΩ. Additionally, the HR of the participant was derived based on the photoplethysmogram measurements that were recorded with a small optical pulse sensor placed on the tip of the participant's index finger.

3.4.4 System setup, tuning, calibration

The execution of the experiment involved numerous stages concerning signal capturing, synchronisation, recording and interfacing with the user. IViewRED and myGaze applications are required for using the heavy and light eye-trackers respectively. These applications provide an interface for calibrating the eye-trackers. A preparatory process is also required for setting up the EEG caps, by ensuring that the skin resistance is kept at low levels (below 10KΩ). A simple application was developed based on BePlus SDK [18] for checking the impedance levels for the Heavyweight configuration. For the EPOC headset we used the official application from Emotiv, Xavier Testbench, ensuring that all electrodes had acceptable level of impedance. Regarding the Shimmer3 device, we used the official Consensys software[19] for programming the device with the latest BtStream firmware and pairing it with our Windows computer via Bluetooth.

As already mentioned, we have used the LSL library for synchronising all the data streams, which essentially is responsible for timestamping each single sample that has been received from the sensors. This process involves using a wrapper application for each sensor in order to communicate with the LSL library. Additionally, events that were specified by each task that was executed by the user were pushed to the LSL network and were also recorded along with sensor data. We used the LabRecorder[20] application in order to record the data from all streams into a single file based on the XDF file format[21].

3.5 Analytics/logging infrastructure

There were two types of user action logging throughout the experiments. First of all a built-in logging functionality from inside the GazeTheWeb Browser saved information to a text file along with a timestamp of each entry. This information included:

- Events such as: mouse over (hover), mouse out, mouse down (click), browser back button, copy action, keyboard strokes, GTW menu screens initiation/actions.
- Training web application specific information: Which page/level has loaded, start or end of a level, success or failure of a level, level specific information about individual game elements,

[17] http://www.shimmersensing.com/products/gsr-optical-pulse-development-kit

[18] http://www.mamem.eu/wp-content/uploads/2016/12/D2.2_Initial_%CE%99ntegration_Optimization_Multi-modal_Sensors_Final.pdf

[19] http://www.shimmersensing.com/support/wireless-sensor-networks-download/

[20] https://github.com/sccn/labstreaminglayer/wiki/LabRecorder.wiki

[21] https://github.com/sccn/xdf

trophies gained.

Secondly an online JSON database through Google Firebase was kept to log the progress of users throughout the training phase. The information saved included:

- User details like, age, gender and nickname that was set up for each user during the registration process to keep anonymity. The training tasks utilized this information to effectively augment the experience with persuasion strategies.
- Level progress information like, completion time, score points, gained trophies.
- Level specific game element information such as, how many times an element appeared or disappeared.
- GTW browser events inside some levels like, mousedown (clicks), mouse out, mouse over (hover).

Based on the Firebase saved user information, after the end of each level of the training tasks, a Leaderboards section was shown with the rankings of users according to score and completion time. More specifically the score for each level was calculated based on the elapsed time from the start of the task until the end, and the amount of errors the user performed (see section 6.1.1.1 for more details).

3.6 Questionnaires

Patient perceived usability and user satisfaction were measured using a standard usability questionnaire (SUS) (Brooke J., 1996) and a user satisfaction questionnaire (QUEST 2.0) (L. Demers et al., 2000). System acceptance was measured by a designated Persuasive Design Questionnaire.

The System Usability Scale (SUS)

Provides a “quick and dirty”, reliable tool for measuring the usability. It consists of a 10-item questionnaire with five response options for respondents; from Strongly agree (5) to Strongly disagree (1). To calculate the total scores, the participant’s scores for each question are converted to a new number, added together and then multiplied by 2.5 to convert the original scores of 0-40 to 0-100. A SUS score above 68 is considered above average and anything below 68 is below average.

The Quebec User Evaluation of Satisfaction with assistive Technology (QUEST)

Is an outcome measurement instrument designed to evaluate a person’s satisfaction with his/her assistive technology device. It can be used with adolescents, adults and elderly persons that use an assistive technology device. The concept of satisfaction as defined in the QUEST 2.0 refers to a person's positive or negative evaluation of those distinct dimensions of the assistive device that are influenced by one's expectations, perceptions, attitudes and personal values. QUEST focus on how satisfied the person is with specific features of the assistive technology device as well as certain characteristics of the services related to the technology (nevertheless this latter part is not relevant to our study and was not used). The QUEST 2.0 was created for assistive technology practitioners and researchers and provides practitioners with a means of collecting satisfaction data that can be used to document the real-life benefits of assistive technology and to justify the need for these devices. The QUEST form displays the scoring of the 12 satisfaction items in two parts: Device (eight items) and Services (four items, not used in this study). The satisfaction items related to the characteristics of the device are: dimensions, weight, adjustments, safety, durability, simplicity of use, comfort and effectiveness. Each item is scored using a 5-point satisfaction scale, with a score of 1 denoting "not satisfied at all" and 5 indicating that the person is "very satisfied". The QUEST 2.0 yields three scores: Device subscale score, Services subscale score (not applicable in this study) and a Total score.

User Acceptance and User evaluation of Persuasive Design Questionnaire

PART I: The user acceptance and evaluation of persuasive design questionnaire – part I, was passed right after the platform training part. In Questions 1-4 the participants were asked to report whether the platform made them feel scared, nervous, unpleasant or unease by indicating whether they agree or disagree with corresponding statements on a scale of 1 to 7. Question 5 asked the participants whether they believe they could operate the platform after they learned to use it alone using the games, or demonstrated how to use it by an instructor by indicating whether they agree or disagree with corresponding statements on a scale of 1 (completely not sure) to 10 (completely sure). Questions 6-14 asked the participants to report on various aspects of the platform such as its ease of use or enjoyability in addition to whether they believe they have enough knowledge to operate it or they believe they had control over it by indicating whether they agree or disagree with corresponding statements on a scale of 1 to 7. Questions 15-17 asked the participants to report on the personalization of the system and whether they believe the games that were used in the training stage motivated them. Finally, question 18 asked the participants whether they would use the system if it were available to them in the future.

PART II: The user acceptance and evaluation of persuasive design questionnaire – part II, was passed right after the dictated tasks part. In Questions 1-3 the participants were asked to report whether the platform could provide better interaction, abilities and output for similar tasks that were tested in the dictated tasks. Questions 4-8 asked the participants to report on whether they found the platform useful, relevant for the type of tasks that were tested in the dictated tasks, whether they found the platform cumbersome, whether they would use the platform if it was available to them in the future and whether they think that most people would learn how to use it fast.

4 Experimental Protocol

4.1 Briefing about the experiment

The aim of the initial phase of the experiment was to inform in depth the volunteers about the project and help the potential participants feel familiar and comfortable with the environment, the MAMEM system and the researchers. Thus, the experimenters introduced themselves as well as the project, the hardware and its objectives and provided the volunteers with non-technical information, concerning the conducted research. This information helped the participant decide if he/she wanted to get involved in the experimental trial and included the risks, benefits and alternatives of the pilot trials along with information describing the nature, the purpose and the benefits of the research process.

The experimenters ensured that the participant was fully aware and capable of fully understanding: i) what the research was about; ii) the objective of the research; iii) for whom the research was being conducted and from where the funding came; iv) what would happen to the results; v) where the results would be published and who would have access to published results and data; vi) what the participant was expected to do if he/she agreed to participate in the pilot trial and how long would the pilot trial last; vii) the constraints concerning the anonymity and confidentiality measures; viii) that he/she was not obliged to participate and his/her right to withdraw from the trial any time without detriment after having agreed to participate; and ix) the project's policy relating to incidental findings.

During the detailed description of the experimental protocol, the experimenters also gave information concerning the equipment used at each stage. Subsequently, the participants had enough time to read and understand thoroughly the consent form, describing all the aforementioned points. After agreeing to participate, the subjects signed the consent form and the first phase of training for the GazeTheWeb

began.

4.2 Interaction Task(s)

4.2.1 Training – GazeTheWeb (Basic, Intermediate, Advanced)

Once the briefing about the experiment was completed, the participants had the opportunity to familiarize with the GazeTheWeb Browser, a custom designed browser that can be operated through the eyes and mind (see also section 3.3). In this part, users only used their eyes to complete a series of tasks that would make them confident about their ability to perform activities without the use of the keyboard or mouse.

Tasks were divided into three main categories (i.e., basic, intermediate and advanced, see also section 3.1) depending on the difficulty level. Basic tasks, aiming on the correct use of the participant's gaze were divided into two sub-categories. Subcategory Basic1 required the user to focus for three seconds on a rhombus-shaped point that randomly appeared on the screen. Once he/she managed to focus on the target, another one appeared on the screen. Subcategory Basic2 was more demanding as the user had a specific time period (4 seconds) at his/her disposal to focus on the rhombus-shaped point (for two seconds) before the next rhombus emerged.

Intermediate tasks included browsing and typing activities and the user was instructed about the use of components such as scroll up/down and copy-paste, but more importantly the use of the virtual keyboard. More specifically, subcategory Intermediate1 was designed to familiarize the user with the use of click and scrolling, with the participant searching for a wizard throughout the screen. Intermediate2 introduced the zoom button and the use of the keyboard as the participant had to answer to 5 multiple choice question by using the virtual keyboard to provide the answer. Finally, in Intermediate3 the user should use the copy-paste module for two longitude and latitude coordinates. More specifically, the user had to select the starting and ending point of the coordinate to be copied and then use the paste button in a text frame.

The last category focused on the use of the interaction elements provided by the GazeTheWeb Browser, within four tasks. During Advanced1 the user had to enable a feature through the settings element and then disable it to complete the task. In Advanced2 a new tab had to be added and the url *"www.mamem.eu"* had to be provided via the virtual keyboard. Advanced3 aimed in the use of auto-completion while typing the phrase "See In The World" with the selection of the appropriate suggestion. The final task of this category (Advanced4) introduced the use of bookmarks, therefore the user had to bookmark a page, move to another tab and then return to the previously bookmarked url.

4.2.2 ErrPs

The distinct neuronal responses that are produced by the human brain during the perception of a mistake are referred as Error Related Potentials (ErrPs) (Kalaganis et al., 2017a; Kalaganis et al., 2017b). Detecting brain commands in a BCI framework is an error-prone procedure, forcing the users in unintentional interaction errors. Thus, ErrP signals could be exploited in order to create a more responsive interface that would enhance the user experience. Herein the experimental procedure is described for an ErrP experiment tailored to a gaze-based (on-screen) keyboard.

Initially, the experimenter shortly introduced the experiment and its objectives to the participants. The objective of this experiment is to measure the brain activation of each participant while typing words and identify the brain signals of unintentional letter typing. The experimenter stressed to the participants that these experiments were conducted in order to explore the brain activity during erroneous actions and not to test their ability in typing tasks. For this, a more challenging keyboard was used which locks faster on

the letters.

The letters were selected after a dwell time of 0.5 seconds and then there was a preview of the selected letter by turning its color to white for 1 second. During the preview time, the participant was advised not to gaze away from the selected letter and not to move to the next one immediately since we are trying to capture their brain activations after each letter selection.

It was made clear to the participants that erroneous actions should be perceived in a letter-by-letter manner and be corrected as soon as possible since checking for errors (e.g., miss-typed letters) at the end of each word or even the whole sentence would not add any value in our research. Moreover, subjects were asked to refrain their movement in order to avoid artifacts (such as eye blinks, jaw clenched etc.) that contaminate the EEG signal.

Five sentences (S1-S5) were shown to the subjects, written clearly with large fonts, in order to be typed using the provided keyboard that operates with an eye-tracker: S1- “My dog barked at the mail carrier”. S2 - “This sentence has only six words”. S3 - “I paid five dollars to buy a ticket”. S4 - “The conference ended at five in the afternoon”. S5 - “The quick brown fox jumps over the lazy dog”.

During the experimental procedure, the ongoing typing sentence should be accessible at any time. In the preliminary feasibility tests we performed, it was found that a convenient place to “pin” each sentence is the top screen bezel so as not to interfere with the on-screen letters. All sentences should be written using lowercase letters. The end of each sentence was followed by short-time breaks. At the beginning of each session, participants typed their name or some other “foo” phrase in order to familiarize with the keyboard settings.

In certain occasions, where the participant encountered great difficulty in writing (e.g., more than 50% of typed letters were errors), the experimenter increased the required dwell time. Moreover it was in experimenter’s discretion to aboard the procedure under certain circumstances (e.g., when the total writing time exceeded 30 minutes).

4.2.3 SMR

Sensorimotor Rhythms (SMR) are brain waves that appear on EEG recordings from areas of the brain that are associated with planning, control, and execution of voluntary movements. In the next paragraphs, a short description of the experimental protocol for elicitation of SMR is provided. Then a software platform related to the protocol, the OpenVibe platform, is shortly described.

The SMR experimental protocol consists of two basic steps. At the first step, the EEG data are acquired without showing any visual feedback to the user (calibration step), while, at the second step feedback is incorporated into the overall procedure (feedback step). In both steps a significant number of EEG trials are acquired for further processing. The trials acquired during the first step are gathered in ordered to train a classifier to be exploited in the subsequent feedback step, while the ones from the latter step are related to the training of the user. Thus, by providing feedback the user is learning to adapt to the system by regulating appropriate brain rhythms.

At the first step (calibration step), the subject tries to imagine the execution of a movement from the left or right hand. Each trial starts with a fixation cross. Some seconds later a visual cue (an arrow pointing either to the left or right) is presented. Afterwards the subjects have to imagine the corresponding hand movement over a period of time, usually 4 to 6 seconds (imaginary period). Each trial is followed by a short break. Also, a randomized short period is added to the break in order to avoid subjects adaptation. At the feedback step, feedback is provided to the subject during the imaginary period with respect to the

classification of their EEG signals. Various types of feedback could be used at this stage (e.g., a bar, a smiley, etc.). Depending on the cue, the subject tries to move accordingly the feedback (usually towards the left or right side) by imagining left or right hand movements, respectively.

In our analysis, two sessions were recorded for each subject at the same day. In the first session, the subject was instructed to execute the real movement while in the second session to imagine the movement. Each session consists of two parts (calibration and feedback steps). The real movement session was applied only to the participants that were able to move their fingers (i.e. all healthy participants and the patients with Parkinson disease). The calibration step consisted of one run, while the feedback step consisted of 4 runs. In each run 20 trials (10 trials for each class, right or left) were collected. The timing events for a trial belonging to the calibration step are provided in Figure. 2, while the timing events of a trial from the feedback step are shown in Figure. 3.

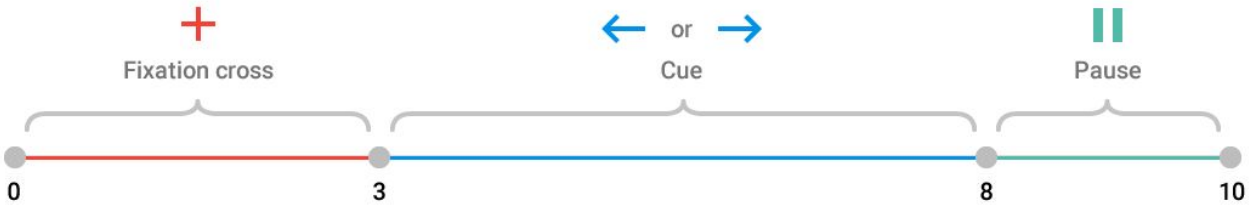

Figure 2: Timing Events of a trial during calibration step.

Figure 3: Timing Events of a trial during feedback step.

The SMR experiment exploits one of the ready-to-use scenarios provided by Open Vibe (Renard et al 2010) that is a free and open-source software platform for the design, test and use of Brain-Computer Interfaces. Designing and operating a BCI experiment with OpenVibe platform requires three distinct steps. In the first step, a training dataset must be recorded for a given subject, while she/he performs specific mental tasks. The second step consists of an offline analysis of these recordings with the goal of finding the best calibration parameters (e.g., optimal features, relevant channels, etc.) for each subject. The last step consists of using the BCI online in a closed loop process. Optionally, iterations can be done on data acquisition and offline training in order to refine the parameters. The online loop (third step) is common to any BCI system and it is composed of six phases: brain activity measurements, preprocessing, feature extraction, classification, translation into a command and feedback. Existing and pre-configured, ready-to-use scenarios are proposed to assist the user of OpenVibe platform. Currently, five complete scenarios are available related to BCI research; hand motor imagery based BCI, self-paced BCI based on foot movements, neurofeedback, real-time visualization of brain activity in 2D/3D, P300-speller. In this study, we used the hand motor imagery based BCI scenario, which allows to use OpenVibe as an interaction peripheral using imagined movements of the left and right hand. This scenario is inspired from the well-known Graz-BCI experimental protocol (Jeunet et al 2016) of the Graz University. All the experiments were designed using the OpenVibe platform as the stimulation source and

feedback initiation (see Appendix 9.3 for more details on the utilized OpenVibe configuration).

4.2.4 Dictated Tasks - GazeTheWeb

The Dictated Tasks was the last part of the experimental protocol and was a simulation of GazeTheWeb browser usage (see section 3.3) in everyday activities. Participants needed to complete four different scenarios (sending an email, editing an image, posting a tweet and watching a YouTube video) composed by several steps that were verbally provided to them by the experimenter. All scenarios were initiated with the three following steps: Tab Overview, Add Tab and selecting the proper link regarding the scenario.

In the first scenario the user ought to reply to an email via Gmail. In order to do so, he/she should sign in using the virtual keyboard with the credentials of an account created for the experiment's purposes. A specific mail from the inbox should be opened and the reply button should be selected. In the email body the phrase "Hello World" should be inserted and hitting the send button would conclude the scenario.

The second scenario included the manipulation of an image file with the use of *picresize*, an online photo editing software[22]. The user was asked to select one of the three sample images to be edited. The next step required the participant to rotate the selected image by selecting the Rotate 90º CCW option, followed by the selection of one special filter effect to be applied on the image. The completion step was to select the button "I'm Done, Resize My Picture!".

The third scenario aimed in the use of Twitter. The participant, using an account made for the experiment's purposes, had to search for the account "MAMEM Project" and follow it. Next, a tweet shall be posted. Users were encouraged to create a tweet with few words stating their experience during the experimental procedure.

Finally, the use of YouTube was simulated in the last scenario and was initiated with the search of the keyword "gazetheweb mamem". The next step was to select one of the videos provided by the searching process. Once the video was initiated, it should be paused and played while focusing in the video (like clicking on the video in order to pause/play it). The final step was to close the last opened tab using the appropriate button.

4.3 Execution Procedure

At the beginning of the experimental procedure, the participant filled in and signed the necessary consent form and was briefed about the incidental findings policy. Also some additional information about the participant was collected, including demographic & clinical data. The participant was checked against inclusion/exclusion criteria for the rest of the experimental process and was classified in a persuasive design group (see section 3.2). The staff explained briefly to the participant the overall project, the purpose and procedure of the trials, and utilized technology. The heavy-weight system was set up and tested. This step included the initialization of the Red N (SMI) eye-tracker, the activation of the GSR sensor, the initialization of the Lab Recorder and all other applications that were required for simultaneous signal collection.

The GTW-Game was initialized and the participant filled-in an online form with his credentials to start the game. The staff ensured that the participant understood the purpose of the game and explained the basic mechanisms of the interface such as scoring, error-counting, time keeping and trophy acquisition. The participant executed all basic, intermediate and advanced tasks of the GTW-Game. The staff was

22 http://www.picresize.com/

always available to intervene in case the participant was frustrated. After finishing each group of tasks in the game, the participant filled-in a corresponding task analysis sheet for all task groups.

After the end of the GTW-Game training tasks, the staff stopped the Lab Recorder and saved the generated files. There was a 5 minute break and after that, the Part-1 Persuasive Design and System Usability questionnaires were filled-in by the participant. The staff proceeded with putting on the participant the EBN Cap, and starting all related recording applications. The participant was also briefed about the ErrPs keyboard experiment.

The participant started the new task, using the ErrPs-designed GTW keyboard, and at first the system collected recordings in resting state using a fixation cross for 3 minutes. After that, the participant was asked by the staff to type a set of predefined sentences (see also section 4.2.2). Another 3 minute resting state phase initiated and then the ErrPs trial concluded with the staff saving all generated files. Subsequently, the SMR experiment launched, without the use of the eye-tracker. It constituted of a 3 minute resting state with a fixation cross, two SMR experiments and another 3 minute resting state phase.

After a 15 minute break, the Dictated tasks trial initiated. It constituted of social media actions that the participant had to perform. More specifically the actions were: i) Sending an e-mail, ii) Editing a photo, iii) Using Twitter to send a tweet, iv) Using YouTube to play/pause a video. The dictated tasks phase ended with the participant filling in the dictated tasks analysis sheet. Finally, the Part 2 of the Persuasive Design Questionnaire was filled-in by the participant.

Subsequently, there was a 50 minute break and, for a selected set of participants, the second part of the trials commenced. The staff switched the experimental configuration to the lightweight one. More specifically, the RedN eye-tracker was replaced with the myGaze eye-tracker and the EBN cap was removed from the participant. The participant then put on the Emotiv EEG cap and the same ErrP experiments took place. There was also 2 phases of 3-minute resting states with a cross fixation as above, at the start and end of the ErrP experiment. The SMR experiment was also performed as above with the Emotiv EEG cap. After the end of the lightweight configuration trials, the participant filled in the Quest 2.0 questionnaire concerning the lightweight configuration. Figure 4 provides a diagrammatic overview of the execution procedure.

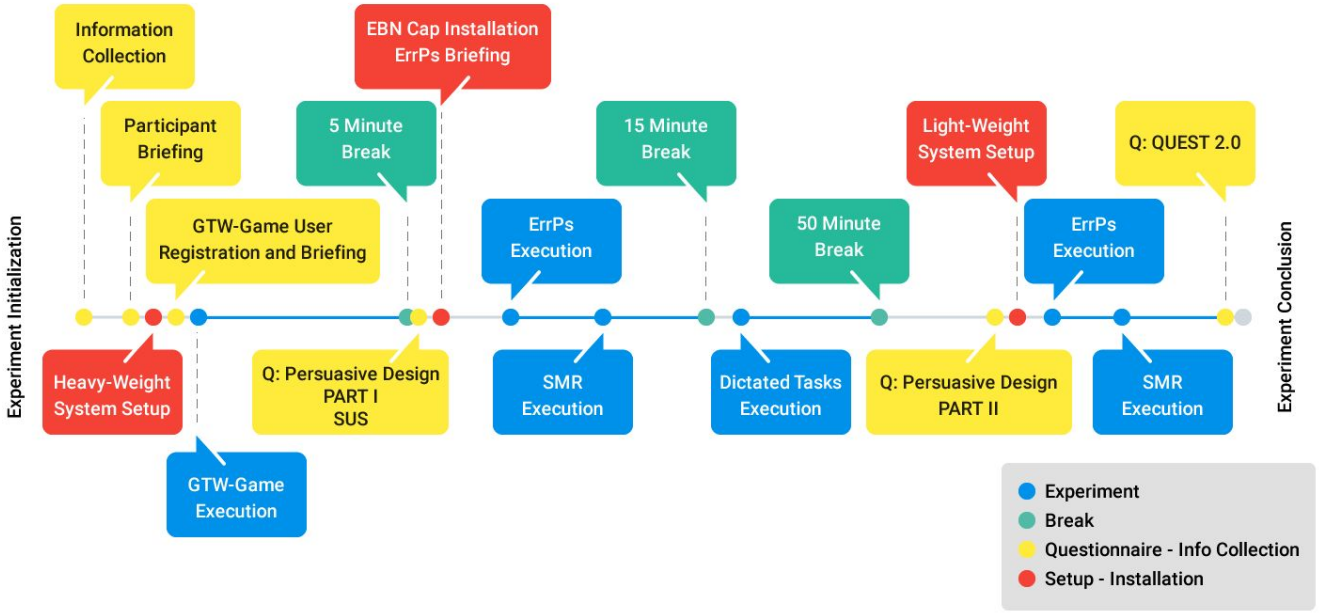

Figure 4: Experimental procedure timeline.

5 Signals and dataset format

The dataset of signals described by this report is divided in three major folders concerning the three experimental sites, i.e., AUTH, MDA, and SHEBA. Inside each aforementioned folder there are subfolders regarding to each subject separately and for each subject .mat files are provided including the recorded data. The latter files correspond to the four different experimental phases plus two for the lightweight configurations, if available. Specifically, the experimental phases regard to: a) the signals recorded during the training process (GazeTheWeb (GTW) training), b) during the ErrP experiment (Heavy and Light-weight configuration), c) during the SMR experiment (Heavy and Light-weight configuration), and d) during the Dictated tasks. Each .mat file is named with the following pattern, [SITE]_[SUBJECTID]_[Experiment], where "SITE" can be either AUTH, MDA or SHEBA, "SUBJECTID" refers to the ID of each subject (See Table 1) and "Experiment" can be either one of the following; a) GTW, b) Heavy_ERRP, c) Light_ERRP, d) Heavy_SMR, e) Light_SMR, and f) DICTATED.

Inside each .mat file lies a struct variable that has been generated from the LabRecorder application. The size of that variable indicates the number of streams that have been recorded during the associated experimental phase. Specifically, six types of streams can be found in each file of the dataset. In particular:

- **Shimmer_[PORT]** stream includes the measurements that have been recorded from the Shimmer device. The data is structured as a 2xN matrix of N samples, in which the first channel of data is the HR measurements in beats per minute and the second channel is the GSR measurements in KΩ. Both data are sampled at 256 Hz and are always equally sized. For the HR values the "PPG to HR" algorithm that is available for the Shimmer device was used, which estimates the HR values based on the RR intervals between two successive heart beats, thus the values are updated by the time a new heartbeat is detected by the sensor. The data quality is affected in some cases due to the loss of contact of the sensors. Loss of contact in the GSR sensors usually appears on the signal as large peaks and in some rare cases as negative values. In HR data, loss of contact can be identified by 0

values followed by a relatively large peak in the data. It is important that both of these cases should be taken into account before passing the data into any processing algorithm.

- **EmotivLSL_EEG** stream refers to the light-weight EEG device that was used by some participants. The data consists of 14 EEG channels (see Appendix 9.1), sampled at 128 Hz. Impedance levels of the sensors were all reported as OK (green color) by the official Emotiv application.
- **EBNeuro_BePlusLTM_[IP_address]** stream consists of the EEG data from the heavy-weight device. The data are structured as a 62xN matrix of N samples (see the Appendix 9.1 for the corresponding electrodes). Impedance levels of the electrodes were kept as low as possible and always below 10KΩ. Due to frequent usage of the cap prior to the experiments, the C3 channel was not working during the experiments that were performed in the sites of AUTH and MDA, so the data in the electrode #25 consists only of noise.
- **iViewXLSL & myGazeLSL** streams consist of eye-tracker data that were captured from the heavy and light-weight eye-trackers, respectively. Both streams are identical in structure, including two channels of gaze data, the first being the gaze coordinates in pixels, along the horizontal axis of the screen and the other one on the vertical. A difference between the data from the two streams is the maximum sampling rate between the two eye-trackers, being 60Hz for the heavy-weight version and 30Hz for the light-weight version. An important thing to note as well, is that in each site, screens with different resolutions were used, so the eye-tracker will be different in terms of scale (i.e. the data range of x,y values will be different for each site). In data produced in AUTH site, the screen resolution was 1680x1050 pixels, in MDA site 1920x1080 pixels and in SHEBA 1366x768. In some rare cases, the eye-tracker stopped working unexpectedly during the experiments, thus some data points might be missing.
- **BrowserOutputStream (Events stream)** is a stream of text markers that are generated from GazeTheWeb during user interaction. This includes key typing, mouse events, web navigation events and also events that are task-specific. A complete list of the events that are logged can be found in the Appendix 9.2.

Each of the aforementioned streams is another struct with three fields: a) info, b) time_series and c) time_stamps. The "info" field provides several information about the stream, including the name and type of the stream which can be used for identifying it (See xdf format specification[23] for more information). The "time_series" field includes the actual data of the stream and the equally-sized "time_stamps" field provide the timestamp of each data point. For the case of the **BrowserOutputStream (**hereafter referred as *events* stream) there are also additional fields that include the nearest sample indices corresponding to the rest of the streams. For instance, an event "X" will have the following additional fields: a) index of the EEG sample that was recorded at the same timestamp as the event, b) similarly for the eye-tracker stream, and c) similarly for the bio-measurements stream. The abovementioned structure of the dataset is illustrated in Figure 5. The full dataset has been made available in (MAMEM Phase I Dataset., 2017). Moreover, appropriate Matlab toolbox for loading and processing the corresponding signals can be downloaded from GitHub platform[24].

23 https://github.com/sccn/xdf/wiki/Specifications

24 https://github.com/MAMEM/eeg-processing-toolbox

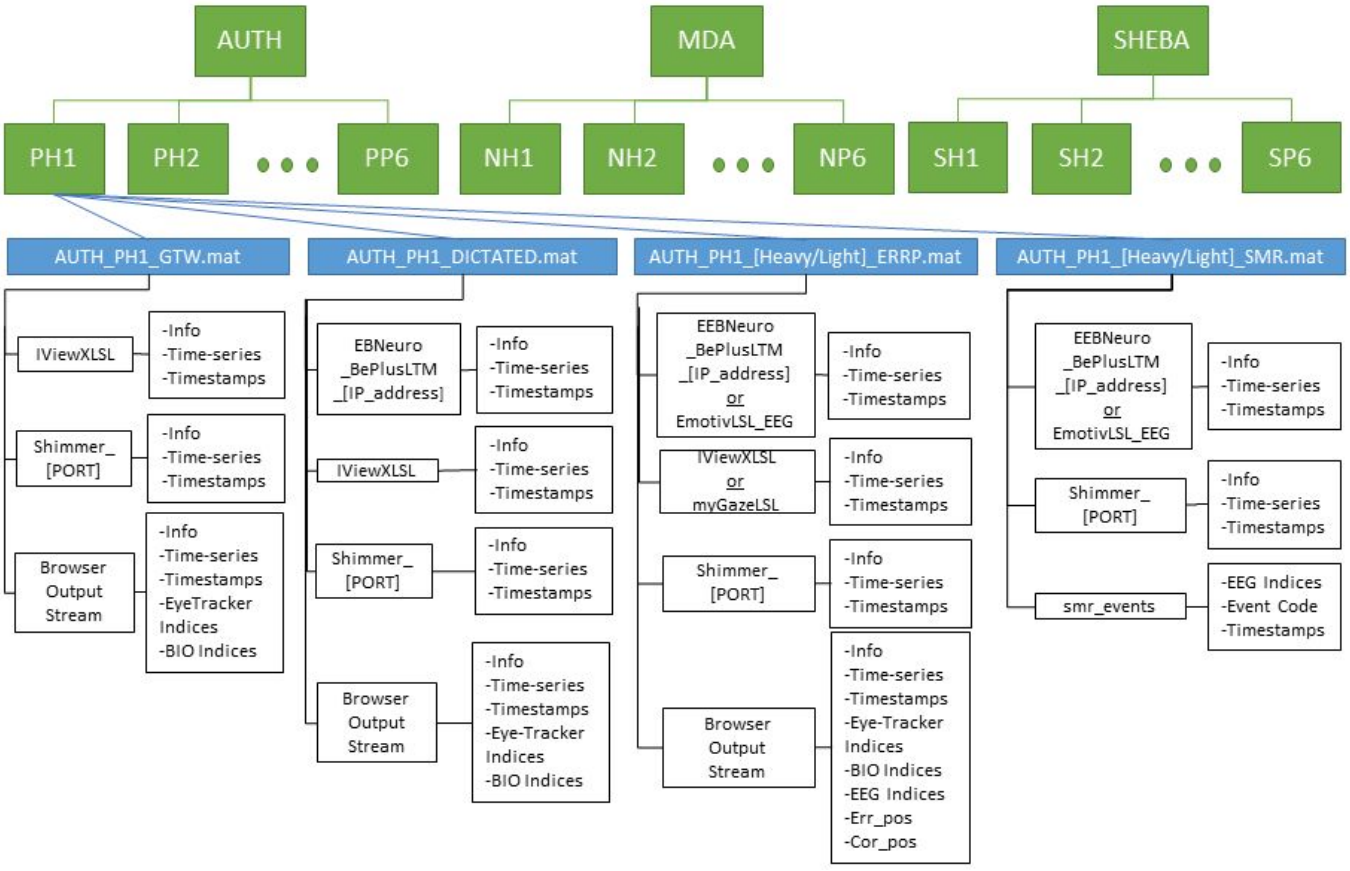

Figure 5: Database structure.

5.1 GazeTheWeb training

The first phase of the experiment includes the GazeTheWeb training, when the participant is introduced to the functionalities of our gaze-controlled web browser. The signals that are recorded during this task are: a) GSR & HR, b) Eye-tracker gaze and c) a stream with events that are associated with the user's behaviour when completing the training levels. The beginning of each subtask is marked by an event such as "*page_load__level_basic_1_start*" and ends with a "*level_complete*" or a "*level_failed*" event, based on whether the user was successful on completing the task or not. Additionally, there are events that log the progress of the user during each task and whether a trophy was awarded. See the Appendix 9.2 for a complete list.

5.2 ErrP experiment

During the ErrP experiments, the participant is also wearing the EEG cap (either the Heavy or Light-weight version), so an EEG stream has been also included. In this task, the keys that were pressed by the user were logged by the GazeTheWeb application. These events have been analysed offline, and for each key press of the user we extracted the information on whether the key press was correct or a mistake based on the sentence that he/she is supposed to write (see Section 4.2.2). Finally, there are also events that mark the beginning and the end of each sentence as well as a resting period which can be used as a baseline signal.

5.3 SMR experiment

The SMR experiment, as specified in Section 4.2.3, was performed with the help of OpenVIBE for recording of eight sessions in which the participants were asked to move or imagine moving their right or left fists. All the sessions that have been recorded by each participant was merged into a single file,

annotated by string markers which indicate the beginning and ending of each session. The total number of sessions may be less than eight for some participants due to their fatigue or incapability of physically moving their hands. Additional events have been added corresponding to what the user is instructed to do, or what elements are appearing on the screen during the experiment.

The events that are included in the merged file are structured differently from the events during different phases. Specifically, a variable named "smr_events" has been included in each file with size 3xN where N is the total number of events. The first row of this matrix denotes the index of the closest EEG sample in terms of time relative to the event, the second row includes the actual event code and the third row includes the timestamp of the closest EEG sample. The timestamp value may also be used for finding the closest sample of the GSR and HR streams. The complete list of the event codes and their description can be found in the Appendix 9.2 under the Event List section.

5.4 Dictated Tasks

In the last phase of the experiment, the signals that are recorded are: a) eye-tracker, b) EEG, c) GSR & HR, and d) events from GazeTheWeb. This session is divided in 4 sub-tasks, namely the tasks of sending an e-mail, editing a photo, posting a tweet and watching a video on youtube. In these tasks, the participant had more freedom regarding the specific steps to perform and for that reason each file might be slightly different than the other in terms of structure. However, it is still possible to break down the whole session in subtasks by assuming that the beginning of each task is marked by the corresponding "Loading URL:" event followed by the appropriate link for each task.

6 Evaluation

6.1 Objective evaluation

6.1.1 Accuracy, time, compose scores

Evaluation of the participant's performance in the training part of the GazeTheWeb arises from three factors, i.e., *time*, *compose score* and *accuracy* which can provide information that can be used for statistical analysis and future improvement. In particular, the first factor taken into consideration is the time needed to complete each task. *Compose score* which depends on the time value, comprise the second factor, and can also contain extra information depending on the task (for a detailed description of the tasks see section 4.2.1). More specifically, compose scores for all Advanced tasks and Intermediate3 only the total time of each task divided by the number of total steps is used to calculate the score, whereas for the remaining tasks extra factors are used. For Basic1 the total number of countsOffs (i.e., A counter that keeps track of how many times the cursor exited an element) is used, for Basic2 the ratio of moles (i.e., the time required to focus on one mole) is used, for Intermediate1 the number of clicks and for Intermediate2 the number of failed efforts. The exact formulas are provided in **Table 9**.

Table 9: GazeTheWeb Training Part: Compose Scores

Task	Compose Score
Basic1	500.000 - [(total_time/2) + (countOff * 50)]
Basic2	total_score + (segment * ratio[mole])
Intermediate1	100.000 - (total_time/2 + (clicks * 50))

Intermediate2	100.000 - (total_time/4 + (fail * 200))
Intermediate3	100.000 - (total_time/4)
Advanced1	100.000 - (total_time/4)
Advanced2	100.000 - (total_time/6)
Advanced3	100.000 - (total_time/6)
Advanced4	100.000 - (total_time/6)

The last factor, *accuracy*, does not include time and is calculated differently for each task:

Basic1&2. The accuracy for the Basic1 task measures how well the users were focusing their gaze on the targets and is reduced for each time their gaze left the target. It is calculated based on the following equation, where *nhits* is the amount of targets that were successfully hit and *nmiss* is the number of times the participant's gaze left the target without completing it.

$$Bas1_{acc} = \frac{nhits}{nhits + nmiss}$$

The accuracy is calculated in a similar manner for the Basic2 task, but in this case, since the targets disappear after a few seconds, the accuracy is also reduced if the user misses a target completely. In this case, *n* includes the amount of markers that had timed-out in addition to the ones that have been hit.

$$Bas2_{acc} = \frac{nhits}{n + nmiss}$$

Intermediate1,2,3. The accuracy for Intermediate1 is based on the number of clicks (*nclicks*) the user made to find the wizard, with the optimal number *oclicks* of clicks being two.

$$Int1_{acc} = \frac{oclicks}{(nclicks - oclicks) + oclicks}$$

The accuracy for Intermediate2 depends on the number of correct answers (*ncorrect*) the user provided in the multiple-choice question (*nquestions*) task.

$$Int2_{acc} = \frac{ncorrect}{nquestions}$$

The accuracy for Intermediate3 measures the number of clicks (*nclicks*) the user pressed during the copy paste procedure. The user has to complete the procedure twice with the optimal number of clicks (*oclicks*) being different in the persuasive and non-persuasive design.

$$Int3_{acc} = \frac{oclicks}{(nclicks - oclicks) + oclicks}$$

Advanced 1,2,3,4. The accuracy for Advanced1 is calculated based on the number of times the user enable the *toggle view* feature (*gazeon*) and the ones the user disabled it (*gazeoff*). The optimal number (*oclicks*) for the enabling and disabling process is considered to be two.

$$Adv1_{acc} = \frac{oclicks}{(gazeon+gazeoff - oclicks) + oclicks}$$

The accuracy for Advanced2 counts the number of keystrokes (*nkeys*) and compares them with the number of the expected keystrokes (*okeys*) the user should make while typing *www.mamem.eu*.

$$Adv2_{acc} = \frac{okeys}{(nkeys - okeys) + okeys}$$

For Advanced3 the number of times the user pressed the exit button (*nclicks*) is measured. The user is

expected to hit this button only once (*oclicks*).

$$Adv3_{acc} = \frac{oclicks}{(nclicks - oclicks) + oclicks}$$

Finally, regarding Advanced4 the number of times the *select bookmark* (*nselect*) and *add bookmark* (*nadd*) buttons are pressed by the user is counted and compared with the number of the times the user was expected to press these two buttons (*oclicks*).

$$Adv4_{acc} = \frac{oclicks}{(nselect+nadd - oclicks) + oclicks}$$

6.1.2 Biosignals (SMR & ErrPs) (Heavy-weight configuration)

The experimental staff was responsible for evaluating the participant's performance in the ErrPs and SMR tasks by providing marks ranging between 1 and 5. Although this part is clearly subjective there is a rationale about each task and grade. Regarding ErrPs the number of mistakes, the ability to recognize them and the completion of sentences were the factors that determine the grade given for each sentence. For the SMR trials, grades are only provided for the feedback sessions and correspond to the percentage of correct trials in the four feedback sessions. A more extensive description is provided in **Table 10**.

Table 10: ErrPs & SMR Evaluation

Grade	ErrPs - Description	SMR - Description
1	Did not complete all sentences/Did not recognize mistakes	50% correct trials
2	Did not recognize mistakes	50-60% correct trials
3	Made a lot mistakes but was able to recognize them	60-70% correct trials
4	Made a few mistakes and corrected them	70-85% correct trials
5	Made almost no mistakes and corrected them	>85% correct trials

6.1.3 Typing speed

Based on the key typing logs that were recorded during the GazeTheWeb training and dictated tasks, we have extracted information about the typing speed of each user during the experiment. Specifically, the tasks that were included during this analysis was the "quiz" game from the training phase, as well as the e-mail, twitter and youtube tasks. The typing speed in characters/second for these four tasks can be found on the CRF file (supplementary material) in columns a) Intermediate Tasks 2 Progression b) E-mail progression c) Social Network- Progression and d) Youtube - Progression, respectively. Each number that is separated by the ';' character corresponds to the typing speed at the given time based on the following formula.

$$S(i) = \frac{1}{t_i}$$

Where S denotes the typing speed in characters/second and t_i is the time that elapsed between typing the i^{th} character and the previous character. In the case that the i^{th} character is the first character, the time that the user opened the keyboard interface is taken into account instead of the previous character.

6.2 Subjective evaluation

6.2.1 System usability (SUS) and user satisfaction questionnaire (QUEST 2.0)

The SUS scores were calculated by assigning a relative score to each item and performing a calculation with their sum[25]. The QUEST 2.0 questionnaire was answered only by those who tested the light weight configuration, half of them were able-bodied participants and half were motor-impaired participants, and the participants were asked to answer in relation to this configuration only. The QUEST 2.0 scores were calculated by averaging the first part of the questionnaire that concerns the different physical and usability aspects of the assistive device. The second part was not filled since it concerns the support and service of the assistive framework which were not relevant in this study.

The average score of QUEST 2.0 for SCI, PD and their matched able-bodied participants was ranging from 4.2 for motor-impaired to 4.4 for able-bodied which is between "*very satisfied*" and "*quite satisfied*". The average user satisfaction score was 3.7 for the able-bodied and 3.1 for the NMD participants, indicating a *"moderately satisfied"* response on a 5 grade Likert scale. The perceived usability (SUS score) for NMD, SCI and all matched able-bodied participants was over 70.0, an above average score (68.0), while for PD was 58.7.

6.2.2 User acceptance and user evaluation of persuasive design - Part I

The user acceptance and evaluation of persuasive design questionnaire – part I, was passed right after the platform training part. Both able-bodied and PD expressed nervousness with able-bodied subjects scoring higher than PD patients. However, general acceptance of persuasive design was gained by both groups. Similarly, reactions to user acceptance questionnaire part I were observed between the groups of patients with NMD and SCI and their matched able-bodied. The results show that SCI participants found the training games less motivating and the platforms' messages less personalized.

6.2.3 User acceptance and user evaluation of persuasive design – part II

The user acceptance and evaluation of persuasive design questionnaire – part II, was passed right after the dictated tasks part. In general, reactions to User Acceptance Questionnaire Part 2, questions 1-3, did not differ between the groups of patients (PD, NMD) and their matched able - bodied participants. Albeit, the results show that SCI participants believed less that using the MAMEM platform can improve and increase their ability to author multimedia.

6.2.4 Spontaneous Feedback/comments

One of the main objectives of the Phase I trials was to receive feedback for the users' experience and suggestions about future improvements/alterations. The virtual keyboard of GazeTheWeb Browser was the subject of discussion among the majority of the participants, with a major difficulty arising with the lack of Greek or Hebrew version of the keyboard. This was an issue mostly for the Parkinson Disease cohort, as participants were older than 50 years old and could not read or write in English. Furthermore, suggestions were made about the space and backspace button and the change of their color, size and even position as it was not easy to separate them from the surrounding buttons (e.g., backspace button was really close to the "1" button and as a result many times this button was selected instead of backspace). The last request made regarding the keyboard was to include the underscore button, that was not included in this version, as in a couple of cases it was included in the provided email account. Some participants found it also difficult to focus and select small fonts (e.g., reply on Gmail) or one out of many options (e.g., special effect selection on picresize). Concerning the SMR part, some participants requested a

[25] https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

change in the feedback provided during the tasks as they believed that it was not representative in some cases. In addition, most of the participants in Sheba Medical Centre indicated that the process of putting on the EEG cap with the gel took too long and was uncomfortable, while removing it was even more unpleasant. The last suggestion was made by the Neuromuscular Disease cohort regarding the EPOC headset placement. Due to the plastic part placed behind the occipital area they could not lean back to their wheelchair and asked whether this part could be placed in the upper part of the head.

6.2.5 Experimenters diary

Besides the comments provided by the users, remarks were also made by the experimental staff in the direction of future improvements. Based on these remarks we decide to separate them into two categories, those that arise from the clinical state of the participants and the those that arise from the implementation of the platform.

The first category refers to signal capturing problems resulting from the clinical state of the participants. In the PD cohort there were two issues, one with ptosis (eyelid) and one with the head of the participant leaning towards the right side. In the SCI cohort, the eyes of one participants were constantly almost closed (ptosis) and one participant was constantly sitting in a leaning position. In all of these cases, the eye-tracker could not identify the eyes easily and several calibration steps were required. In the NMD cohort two users were constantly moving their upper body in their effort to breathe, creating problems with the eye-tracker that could not always identify their eyes, as well as the EEG recording that was polluted with noisy. Moreover, some of the NMD participants had small fingers that made the GSR adjustment nearly impossible. In addition, when both PD and NMD participants used glasses it was difficult to reach sufficient levels of accuracy in the calibration procedure. Finally, in PD and NMD cohorts there were cases that the EPOC headcap was moving during the experiment and the experimenters had to re-place it. As a result, we cannot be sure that the cup was placed in the exact same position.

The second category includes comments about the platform's components/tasks that users could not use/complete easily. It was difficult for the majority of the participants to focus on the upper left corner of the screen, as well as the rhombus-shaped point in the Basic tasks. The response time of space and backspace button was different from the rest of the buttons making it difficult for the participants to realize whether they were pressed or not. The decision time on the keyboard is fixed (1 second) and there were cases that this time was not enough or was too much for some individuals. Making this time adjustable could make the user experience a lot better.

6.3 Case Reported Files

As in any study involving patients, we collected specific clinical and demographic information for each participant for both motor-impaired and able-bodied ones. In addition, we also collected the different scores of participant's performance in each task and other relevant information of the experimental procedure. In addition, in case report files (CRF, Supplementary Material) there is detailed report of the symptoms, signs, diagnosis, treatment, and demographics (age, sex, education, marital status etc) of each individual patient. All this information has been aggregated for all participants in a single excel file, delivered as part of the dataset. The integration of all these characteristics, of both clinical remarks and performance indicators for every participant is deemed of particular value and can add to the understanding of the performance achieved by each participant in executing the experimental protocol, as well as how each clinical condition or treatment of the disease can influence this performance. This can allow researchers to conduct multiple regression analysis or any other statistical test or apply correlation

analysis to find how specific variables are correlated. In general, this excel file represents a relevant, timely, and important document of the study.

7 Conclusions

In this report we have presented a dataset concerning human computer interaction under a novel framework of authoring and editing multimedia content. To our knowledge this is the first dataset incorporating multimodal signals from many subjects under the aforementioned frameworks and consists of an ideal testbed for development and evaluation of novel human computer interaction algorithms and methodologies. The dissemination of the dataset and its use to test human computer interaction algorithms will further justify its reliability and will pave the way for more accurate and practical systems for helping people with motor impairments to integrate back to society.

8 References

Adedoyin-Olowe, M., Gaber, M. M., & Stahl, F. (2014) "A survey of data mining techniques for social network analysis", Journal of Data Mining & Digital Humanities.

Arroyo, S., Lesser, RP., Gordon, B, Uematsu, S, Jackson, D, Webber, R, "Functional significance of the mu rhythm of human cortex: an electrophysiologic study with subdural electrodes". *Electroencephalography and Clinical Neurophysiology*. **87** (3): 76–87, 1993.

Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, eds. Usability Evaluation in Industry. London, UK: Taylor & Francis; 1996:189-194.

Demers L, Weiss-Lambrou R, Ska B. Quebec user Evaluation of Satifaction with assistive Technology. QUEST version 2.0. An outcome measure for assistive technology devices. Institute for Matching Person and Technology; New York: Webster; 2000.

Fountoukidou S., Ham J., Ruijten P., and Matzat U., Using personalized persuasive strategies to increase acceptance and use of HCI technology, Adjunct Proceedings of the 11th International Conference on Persuasive Technology, Salzburg-Austria, April 2016

Fountoukidou S., Ham J., Midden C., and Matzat U., Using Tailoring to Increase the Effectiveness of a Persuasive Game-Based Training for Novel Technologies, Proceedings of the Personalization in Persuasive Technology Workshop, Persuasive Technology 2017, Amsterdam, The Netherlands, April 2017

Handy, Todd C. *Event-related potentials: A methods handbook*. MIT press, 2005.

Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. J Neurol Neurosurg Psychiatry 1992; 55: 181–84, 1992.

Jeunet C., Jahanpour E., and Lotte F.. Why standard brain-computer interface (bci) training protocols should be changed: an experimental study. Journal of Neural Engineering, 13(3):036024, 2016.

Kalaganis F., Chatzilari E., Georgiadis K., Nikolopoulos S., Laskaris N. and Kompatsiaris I., An Error Aware SSVEP-based BCI, 30th IEEE International Symposium on Computer-based Medical Systems, Special Track on Multimodal Interfaces for Natural Human Computer Interaction: Theory and Applications, IEEE CBMS 2017, June 22-24, 2017a Thessaloniki - Greece

Kalaganis, F., Chatzilari, E., Nikolopoulos, S., Laskaris, N., & Kompatsiaris, I., A Collaborative Representation Approach to Detecting Error-Related Potentials in SSVEP-BCIs. In Proceedings of the 2017 ACM on Multimedia Conference, to appear, October, 2017b.

Kirshblum, S.C., Burns, S.P., Biering-Sorensen F, Donovan W, Graves DE, Jha A, Johansen, M., Jones, L., Krassioukov, A., Mulcahey,M.J., Schmidt-Read, M., and Waring, W. (2011). International standards for neurological classification of spinal cord injury (revised 2011). J. Spinal. Cord Med. 34, 535–546.

Kumar C., Menges R., Staab S., Eye-Controlled Interfaces for Multimedia Interaction, IEEE MultiMedia, 23.4: 6-13, 2016.

Kumar C., Menges R., Mueller D., Staab S., Chromium based Framework to Include Gaze Interaction in Web Browser, (honourable mention) Proceedings of the 26th International Conference Companion on World Wide Web, 219-224, 2017.

MAMEM Phase I Dataset., Dataset for multimodal human-computer interaction using biosignals and eye tracking information, July 2017, doi: 10.6084/m9.figshare.5231053

Oikonomou V., Georgiadis K., Liaros G., Nikolopoulos S. and Kompatsiaris I., A comparison study on EEG signal processing techniques using motor imagery EEG data, 30th IEEE International Symposium on Computer-based Medical Systems, Special Track on Multimodal Interfaces for Natural Human Computer Interaction: Theory and Applications, IEEE CBMS 2017, June 22-24, 2017 Thessaloniki - Greece

Renard Y., Lotte F., Gibert G., Congedo M. , Maby E., Delannoy V., Bertrand O., and Lcuyer A.. Openvibe: An open-source software platform to design, test, and use braincomputer interfaces in real and virtual environments. Presence: Teleoperators and Virtual Environments, 19(1):3553, 2010.

9 Appendix

9.1 BEPlusLTM and EPOC Channel List

The heavyweight configuration of the EEG system consists of the BePlusLTM Bioelectric Signal Amplifier (64 channels) and an EEG prewired headcap in elastic fabric. The headcap has 61+2 electrodes placed according to the ICNS 10-20 system. The electrodes with even numbers are placed on the left hemisphere, whereas the odd ones on the right. A complete list with all electrodes and the corresponding labels of the amplifier are provided on Table A1 .

Table A1 Mapping for BePlusLTM channel indexes and electrode names

AF7	1	FC3	17	CP3	33	PO3	49
AF3	2	FC1	18	CP1	34	O1	50
FP1	3	FC2	19	CP2	35	O2	51
FP2	4	FC4	20	CP4	36	PO4	52
AF4	5	FC6	21	CP6	37	PO8	53
AF8	6	FT8	22	TP8	38	OZ	54
F7	7	T3	23	T5	39	AFZ	55
F5	8	C5	24	P5	40	FZ	56
F3	9	C3	25	P3	41	FCZ	57
F1	10	C1	26	P1	42	CZ	58
F2	11	C2	27	P2	43	CPZ	59
F4	12	C4	28	P4	44	PZ	60
F6	13	C6	29	P6	45	POZ	61
F8	14	T4	30	T6	46	SERVICE	62
FT7	15	TP7	31	FPZ	47		
FC5	16	CP5	32	PO7	48		

The lightweight configuration of the EEG system consists of a wireless headcap, connected via bluetooth, with 14+2(reference) electrodes placed according to the ICNS 10-20 system. The electrode numbers correspond the left and right hemisphere the way they do in the heavyweight configuration.

Table A2 Emotiv EPOC Channel list

AF3	1
F7	2
F3	3
FC5	4
T7	5
P7	6
O1	7
O2	8
P8	9
T8	10
FC6	11

F4	12
F8	13
AF4	14

9.2 Event List

The complete list of the event markers that have been recorded during our experiments can be found in Table A3 below. These events can be used for marking and segmenting the EEG, eye-tracking, GSR or HR signals. In the first section named “GazeTheWeb - Training & Dictated Tasks”, the events that are common for these two tasks are described. Different sections of the table present the events that are unique for the tasks of GazeTheWeb - Training, ErrP experiment and SMR experiment.

Table A3

Event key	Description
GazeTheWeb - Training & Dictated Tasks	
Click performed	
Close Tab Overview	
Close URL input	
Close Tab Overview	
Display bookmarks	
Edit URL	
GAZE_SELECTED_KEY_[KEY]	Specified which key was pressed
Go back	
Inputting Text: [TEXT]	User typed a text and hit the OK button
Loading URL: [URL]	
Open Tab Overview	
Open bookmark: [BOOKMARK]	
Open new tab	
Submitting Text: [TEXT]	User typed a text and hit the submit button
Text input started	
URL input done	
GazeTheWeb - Training	
page_load__overview	Overview page loaded.
page_load__level_intro	Level intro loaded.
page_load__level_intro_countdown_start	The countdown to start the level has begun.
page_load__level_basic_1_start	BASIC-1 has started.
page_load__level_basic_2_start	BASIC-2 has started.
page_load__level_int_1_start	INT-1 has started.
page_load__level_int_2_start	INT-2 has started.
page_load__level_int_3_start	INT-3 has started.
page_load__level_adv_1_start	ADV-1 has started.

page_load__level_adv_2_start	ADV-2 has started.
page_load__level_adv_3_start	ADV-3 has started.
page_load__level_adv_4_start	ADV-4 has started.
page_load__level_results	Open results page.
page_load__level_failed_text	Openg page that tells user to replay the level.
page_load__level_replay	Initiated level load for replaying.
page_load__next_level	Initiated action to go to next level.
level_basic_1__new_marker	New marker added to BASIC-1
level_basic_1__new_markers_batch	Based on evaluation, another batch of markers was added.
level_basic_2__mole_start	Mole appeared.
level_basic_2__mole_end	Mole dissapeared. Time's up.
level_basic_2__moles_new_batch	Add new batch of moles based on eval.
level_basic_2__mole_hit	Mole HIT.
level_int_1__tower_instance	Got inside the tower.
level_int_1__cave_instance	Got inside the cave.
level_int_2__question_asked_easy	Loaded easy question.
level_int_2__question_asked_med	Loaded medium question.
level_int_2__question_asked_hard	Loaded hard question.
level_int_2__answer_correct	Answered correctly.
level_int_2__answer_wrong	Didn't answer correct.
level_int_3__map_on	Clicked on globe to show the map.
level_int_3__coords_submit	Coordinates submitted.
level_int_3__coords_correct	Coordinates submitted are correct.
level_int_3__coords_wrong	Coordinates submitted are wrong.
level_adv_1__msg_settings	Opened GTW Settings View.
level_adv_1__msg_general	Opened GTW General Settings View.
level_adv_1__msg_gaze_on	Activated Gaze mode.
level_adv_1__msg_gaze_off	Deactivated Gaze mode.
level_adv_1__msg_close	Closed any instance of GTW Interface.
level_adv_2__msg_tabs	Opened GTW Tabs View.
level_adv_2__msg_edit	Started editing a URL.
level_adv_2__msg_keystroke	Keys are pressed on the GTW Keyboard.
level_adv_2__msg_close	Closed any instance of GTW Interface.
level_adv_3__msg_close	Closed any instance of GTW Interface.
level_adv_4__msg_tabs	Opened GTW Tabs View.
level_adv_4__msg_bookmark_add	Added current page as Bookmark.
level_adv_4__msg_new_tab	Added new GTW Tab.
level_adv_4__msg_bookmarks	Accessed the Bookmarks View of GTW.
level_adv_4__msg_open_bookmark	Loaded the saved bookmark.

level_adv_4__msg_tabs_again	Returned to GTW Tabs View.
level_adv_4__msg_tab0	Returned to First tab of GTW.
trophy	Got a trophy.
level_complete	The level is complete.
level_failed	The level has failed.
event__mouse_over	The cursor is on top of an element.
event__mouse_out	The cursor is now out of bounds of the previous element.
event__mouse_down	The hand cursor action (left click) has been fired.
event__browser_back	The BACK button of the browser has been pressed.
event__copy	Pasted text.
SMR Experiment	
101	Start of 1st session (real hand movement)
102	Start of 2nd session (real movement with feedback)
103	Start of 3rd session (imaginary hand movement 1)
104	Start of 4th session (imaginary hand movement 2)
105	Start of 5th session (imaginary hand movement with feedback 1)
106	Start of 6th session (imaginary hand movement with feedback 2)
107	Start of 7th session (imaginary hand movement with feedback 3)
108	Start of 8th session (imaginary hand movement with feedback 4)
201	End of 1st session (real hand movement)
202	End of 2nd session (real movement with feedback)
203	End of 3rd session (imaginary hand movement 1)
204	End of 4th session (imaginary hand movement 2)
205	End of 5th session (imaginary hand movement with feedback 1)
206	End of 6th session (imaginary hand movement with feedback 2)
207	End of 7th session (imaginary hand movement with feedback 3)
208	End of 8th session (imaginary hand movement with feedback 4)
768	Start of Trial
769	Left Arrow presented on the screen
770	Right Arrow presented on the screen
781	Feedback appeared on the screen
786	Cross appeared on the screen
800	End of Trial

32775	Baseline signal started
32776	Baseline signal stopped
ERRP Experiment	
Begin_Rest_1	Start of baseline signal recording 1
Begin_Sentence_1	Begin typing sentence 1
Begin_Sentence_2	Begin typing sentence 2
Begin_Sentence_3	Begin typing sentence 3
Begin_Sentence_4	Begin typing sentence 4
Begin_Sentence_5	Begin typing sentence 5
Begin_Rest_2	Start of baseline signal recording 2
End_Rest_1	End of baseline signal recording 1
End_Sentence_1	End typing sentence 1
End_Sentence_2	End typing sentence 2
End_Sentence_3	End typing sentence 3
End_Sentence_4	End typing sentence 4
End_Sentence_5	End typing sentence 5
End_Rest_2	End of baseline signal recording 2
GAZE_SELECTED_KEY_[KEY]	Specifies which key was typed by the user
Key-Selection-Duration to 0.5 seconds	Key dwell time set to 0.5 seconds (fast mode)
Key-Selection-Duration to 1.0 seconds	Key dwell time set to 1.0 seconds (slow mode)
-1	Typed the wrong character
1	Typed the correct character

9.3 OpenVibe Configuration

We have designed the experiment similarly to the example scenario “motor-imagery-csp” that is provided by OpenVIBE. Four different scenarios have been created which include the following steps;

The **Signal - Acquisition** step provides the interface that indicates to the user what kind of movement to perform and when. The “Graz Motor Imagery BCI Stimulator” OpenVIBE box has been used for the interface, configured for producing 10 trials for each class (right/left). This interface starts with a black screen during which a short baseline signal is recorded. Then, an arrow with direction pointing left or right is displayed, indicating the beginning of a trial and the direction of the movement that the user should perform (left/right). The arrow stays on the screen for 5 seconds after which a small pause of 2 seconds is followed and the user rests. The process is repeated until 10 trials for each class are recorded. The whole step is executed 3 times for each participant, twice for imaginary movement sessions and once for a real movement session (only for capable participants). The sessions that are produced by this step are then used to calculate the CSP parameters and train the classifier.

The **CSP - Training** step uses the “CSP Spatial Filter Trainer” OpenVIBE box in order to calculate the parameters of the CSP. The data which have resulted from the previous step are used for this step. Firstly, depending on the configuration, i.e., heavy or light, we select the appropriate channels. For the Heavyweight configuration we selected the electrodes over the FC, C and CP area and for the Lightweight configuration we selected all of them. Next, the signals are segmented into trials that

represent a 4 second window, half a second after the arrow appears on the screen. These trials are then passed to the “CSP Spatial Filter Trainer” box which generates the CSP filter parameters and saves them in order to be used in the next step. The dimension of the CSP filter was set to 6.

In the **Classifier - Training** step, again the signals that were recorded during the Signal - Acquisition step are used to train an LDA classifier. First, we select the appropriate electrodes based on each configuration and then we segment the signals into trials for each class, both in a similar manner as the previous step. Afterwards, we apply a further segmentation to each trial with a small moving window of 1 second length and a step of 1/16 second. For each of these small windows we extract their power and we calculate their logarithms which are then used as features. The features are passed on the “Classifier trainer” OpenVIBE box, configured to use the LDA classifier. Finally, the trained classifier is saved in a file in order to be used in the final step.

During the **Online Session**, the participants are playing a game of handball, in which they are asked to try to move a ball with their mind and score a goal. Similar to the “Signal - Acquisition” sessions, the participants are asked to perform a real or imaginary movement of their hand when an arrow appears on the screen. This time, however, a feedback bar is displayed on the screen informing them on how well they are performing the movement. This bar starts from the middle of the screen and spans along the horizontal axis all the way up to the right edge or the left edge of the screen. The direction of the feedback bar is based on the output of the classifier that is running on the background and the length of the bar is based on its confidence score. The feedback output is refreshed with a rate of 16Hz during its duration of 4 seconds after the arrow appears on the screen. Every time the feedback is refreshed, the last second of the user’s EEG signal is captured and we extract features from it with the same methods that have been described in the previous step. These features are then passed to the trained LDA classifier and the result of the classifier is visualized to the user via the feedback bar. At the same time, based on their performance, a ball moves from the center of the court towards the goals that are located left or right of the screen. If the participants were able to maintain the correct movement for a sufficient period of time, a goal is scored. In each session the participant is asked to score a total of 20 goals, 10 towards the left post and 10 on the right. The online sessions are executed five times for each participant, once using real movement if the participant is able to, and four times using imaginary movement.