Wireless Communications and Mobile Computing

Mobile Assistive Technologies

Lead Guest Editor: Simone Spagnol
Guest Editors: Ádam Csapó, Evdokimos Konstantinidis, and Kyriaki Kalimeri

Mobile Assistive Technologies

Wireless Communications and Mobile Computing

Mobile Assistive Technologies

Lead Guest Editor: Simone Spagnol
Guest Editors: Ádam Csapó, Evdokimos Konstantinidis, and Kyriaki Kalimeri

Editorial Board

Contents

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 8617892, 2 pages
https://doi.org/10.1155/2018/8617892

Editorial

Mobile Assistive Technologies

Simone Spagnol,[1] **Ádam Csapó**,[2]
Evdokimos I. Konstantinidis,[3] **and Kyriaki Kalimeri**[4]

[1]*Department of Information Engineering, University of Padova, 35131 Padova, Italy*
[2]*Department of Informatics, Széchenyi István University, Győr 9026, Hungary*
[3]*Lab of Medical Physics, Medical School, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece*
[4]*Data Science Laboratory, ISI Foundation, 10126 Torino, Italy*

Correspondence should be addressed to Simone Spagnol; spagnols@hi.is

Received 26 March 2018; Accepted 26 March 2018; Published 28 May 2018

Over the past few years, assistive technologies have achieved important milestones in enhancing the quality of everyday life, in terms of both autonomy and well-being, of individuals in need of assistance and care. The convergence process behind ICT in the past decades, mainly from the point of view of technology, but also in terms of management and regulation has led to the widespread availability of digital tools and rich data sources as building blocks for effective solutions towards assistive technologies. Today the promise of technologies that allow individuals to benefit from portable and discrete aids delivered through mobile devices seems more relevant than ever before. Ranging from assistive home automation (domotics) solutions to applications that repurpose smartphones into electronic assistive devices, breakthrough technologies are allowing users with physical and/or mental disabilities to accomplish everyday tasks, as well as to learn and develop social skills more easily.

The goal of this special issue was to create a multidisciplinary forum of discussion on recent advances within the field of mobile technology to support individuals of all ages in need of assistance and care. The issue is comprised of five papers, each focusing on some aspect of rehabilitation, or on ways of mitigating the effects of disabilities through a unique combination of a variety of sensory and display technologies. The technologies of BCI, virtual reality, and auditory displays are especially well represented in the special issue.

The paper by A. Athanasiou et al. (entitled "Wireless Brain-Robot Interface: User Perception and Performance Assessment of Spinal Cord Injury Patients") presents a pair of 8-DoF anthropomorphic robotic arms that can be controlled through kinesthetic motor imagery using the commercially available Emotiv EPOC BCI device. Through an extensive set of user experiments, the authors provide a detailed assessment of how parameters such as age, independence, psychometric evaluation results, and neurological condition (i.e., among both spinal cord injury patients and healthy individuals) influences users' ability to control the system and their perception of the system. Based on the results, the authors conclude that BCI systems enabling users to effectively control multiple degrees of freedom in real-time may well be within short-term reach.

The paper by R. B. Lupu et al. (entitled "BCI and FES Based Therapy for Stroke Rehabilitation Using VR Facilities") also centers around BCI technology but combines it with virtual reality with a view towards supporting the rehabilitation of stroke patients. The paper introduces a system called TRAVEE, which makes combined use of an array of stimulation devices (e.g., functional electrical stimulation) to help the execution of the rehabilitation tasks, as well as monitoring devices (via BCI and electrooculography) and visual feedback (through virtual therapist) to provide users with real-time guidance. Based on a set of evaluations, the authors confirm that this closed-loop approach of connecting action, stimulation, and real-time feedback can go a long way in supporting the rehabilitation process, while at the same time alleviating the need for the continuous availability of trained therapists.

The paper by J. Gomez et al. (entitled "Using Smartphones to Assist People with Down Syndrome in Their Labour Training and Integration: A Case Study") presents a smartphone application called AssisT-Task, which was developed by the authors to help individuals with cognitive

impairments to establish effective learning methods. The application is structured such that caregivers can define tasks that consist of sequences of subtasks and describe each subtask using text and visual figures which can then be retrieved through context-specific QR codes. A detailed analysis provided in the paper highlights the kinds of tasks and types of disabilities with respect to which the application provides the most benefits, and the authors conclude that, based on the evidence, the application may be most suitable to people with Down syndrome who are being trained to get a job.

The paper by by C. Saitis et al. (entitled "Cognitive Load Assessment from EEG and Peripheral Biosignals for the Design of Visually Impaired Mobility Aids") aims to develop an approach that is capable of real-time automated cognitive load analysis for visually impaired people. To that end, the paper investigates the ways in which different indoor and outdoor scenarios can be linked to certain characteristics of EEG sensor data (in particular, a slightly modified version of the ERD/ERS index, which is a well-established measure of cognitive load) in test subjects with different degrees of visual impairment. At the same time, a random forest based model is proposed that uses the measurements of electrodermal activity and blood volume pulse to predict discrete categories based on the ERD/ERS indices. The success of the model suggests that it may be possible to accurately gauge cognitive load using widely available, unencumbered sensors, while the scenario-dependent investigation in the first part of the paper highlights the fact that environmental and subject-specific factors may serve as useful a priori information in the development of robust solutions for electronic travel aids (ETAs).

Finally, the paper by S. Spagnol et al. (entitled "Current Use and Future Perspectives of Spatial Audio Technologies in Electronic Travel Aids") focuses on a wide range of aspects relevant to the implementation of spatialized audio feedback in ETAs for the visually impaired. The paper provides a detailed overview on the acoustic and environmental parameters which influence the spatial perception of auditory stimuli, as well as the software and hardware aspects of spatial sound production, and the ways in which those approaches are used in currently available ETAs. In terms of software, special focus is given to the various ways in which head-related transfer functions (HRTFs) can be described—that is, through generic models, personalized models, structurally motivated models, and analytical or numerical means. In terms of hardware, headsets that present sound without occluding the ears (allowing visually impaired users who rely on input from the physical environment to perceive it at the same time), such as bone conduction headsets and active transparent headsets, are highlighted as being most relevant to ETAs.

Acknowledgments

As guest editors, we would like to thank all the authors and reviewers whose contributions have made this special issue possible. It is our hope that the selection of works contained in the issue will serve as inspiration for further research and development in the area of mobile assistive technologies.

Simone Spagnol
Ádam Csapó
Evdokimos I. Konstantinidis
Kyriaki Kalimeri

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 4798359, 8 pages
https://doi.org/10.1155/2018/4798359

Research Article

BCI and FES Based Therapy for Stroke Rehabilitation Using VR Facilities

Robert Gabriel Lupu,[1] **Danut Constantin Irimia**,[2] **Florina Ungureanu**,[1] **Marian Silviu Poboroniuc**,[2] **and Alin Moldoveanu**[3]

[1]*Computer Engineering Department, "Gheorghe Asachi" Technical University of Iasi, Iasi, Romania*
[2]*EUEDIA Department, "Gheorghe Asachi" Technical University of Iasi, Iasi, Romania*
[3]*Computer Engineering Department, Politehnica University of Bucharest, Bucharest, Romania*

Correspondence should be addressed to Robert Gabriel Lupu; lupu.robert@gmail.com

Received 22 September 2017; Accepted 14 February 2018; Published 5 April 2018

Academic Editor: Evdokimos I. Konstantinidis

In recent years, the assistive technologies and stroke rehabilitation methods have been empowered by the use of virtual reality environments and the facilities offered by brain computer interface systems and functional electrical stimulators. In this paper, a therapy system for stroke rehabilitation based on these revolutionary techniques is presented. Using a virtual reality Oculus Rift device, the proposed system ushers the patient in a virtual scenario where a virtual therapist coordinates the exercises aimed at restoring brain function. The electrical stimulator helps the patient to perform rehabilitation exercises and the brain computer interface system and an electrooculography device are used to determine if the exercises are executed properly. Laboratory tests on healthy people led to system validation from technical point of view. The clinical tests are in progress, but the preliminary results of the clinical tests have highlighted the good satisfaction degree of patients, the quick accommodation with the proposed therapy, and rapid progress for each user rehabilitation.

1. Introduction

The worldwide statistics reported by World Health Organization highlight that stroke is the third leading cause of death and about 15 million people suffer stroke worldwide each year [1]. Of these, 5 million are permanently disabled needing long time assistance and only 5 million are considered socially integrated after recovering. Recovering from a stroke is a difficult and long process that requires patience, commitment, and access to various assistive technologies and special devices. Rehabilitation is an important part of recovering and helps the patient to keep abilities or gain back lost abilities in order to become more independent. Taking into account the depression installed after stroke, it is very important for a patient to benefit from an efficient and fast rehabilitation program followed by a quick return to community living [2]. In the last decade, many research groups are focused on motor, cognitive, or speech recovery after stroke like Stroke Centers from Johns Hopkins Institute [3], ENIGMA-Stroke Recovery [4], or StrokeBack Consortium funded by European Union's Seventh Framework Programme [5]. Important ICT companies bring a major contribution to the development of technologies and equipment that can be integrated into rehabilitation systems. For example, *Stroke Recovery with Kinect* is a research project to build an interactive and home-rehabilitation system for motor recovery after a stroke based on Microsoft Kinect technology [6].

In the last years, the virtual reality (VR) applications received a boost in development due to VR headset prices that dropped below $1000, allowing them to become a mass-market product [7]. The VR was and still is especially used for military training or video games to provide some sense of realism and interaction with the virtual environment to its users [8]. Now it attracts more and more the interest of physicians and therapist which are exploring the potential of VR headset and augmented reality (AR) to improve the neuroplasticity of the brain, to be used in neurorehabilitation

and treatment of motor/mental disorders [9]. However, considering the diversity of interventions and methods used, there is no evidence that VR therapy alone can be efficacious compared with other traditional therapies for a particular type of impairment [10]. This does not mean that the potential of VR was overestimated and the results are not the ones that were expected. The VR therapy must be complemented with other forms of rehabilitation technologies like robotic therapy, brain computer interface (BCI) and functional electrical stimulation (FES) therapy, and nevertheless traditional therapy to provide a more targeted approach [11].

*Saebo*VR is a virtual rehabilitation system exclusively focusing on activities of daily living and uses a virtual assistant that appears on the screen to educate and facilitate performance by providing real-time feedback [12]. The neurotechnology company MindMaze has introduced MindMotion PRO, a 3D virtual environment therapy for upper limb neurorehabilitation incorporating virtual reality-based physical and cognitive exercise games into stroke rehabilitation programs [13]. At New York Dynamic Neuromuscular Rehabilitation, the CAREN (Computer Assisted Rehabilitation Environment) based on VR is currently used to treat patients poststroke and postbrains injuries [14]. EVREST Multicentre has achieved remarkable results regarding the use of VR exercises in stroke rehabilitation [15].

Motor imagery (MI) is a technique used in poststroke rehabilitation for a long time ago. One of its major problems was that there was not an objective method to determine whether the user is performing the expected movement imagination. MI-based BCIs can quantify the motor imagery and output signals that can be used for controlling an external device such as a wheelchair, neuroprosthesis, or computer. The FES therapy combined with MI-based BCI became a promising technique for stroke rehabilitation. Instead of providing communication, in this case, MI is used to induce closed-loop feedback within conventional poststroke rehabilitation therapy. This approach is called paired stimulation (PS) due to the fact that it pairs each user's motor imagery with stimulation and feedback, such as activation of a functional electrical stimulator (FES), avatar movement, and/or auditory feedback [16]. Recent research from many groups showed that MI can be recorded in the clinical environment from patients and used to control real-time feedback and at the same time, they support the hypothesis that PS could improve the rehabilitation therapy outcome [17–21].

In a recent study, Irimia et al. [22] have proved the efficacy of combining motor imagery, bar feedback, and real hand movements by testing a system combining a MI-based BCI and a neurostimulator on three stroke patients. In every session, the patients had to imagine 120 left-hand and 120 right-hand movements. The visual feedback was provided in form of an extending bar on the screen. During the trials where the correct imagination was classified, the FES was activated in order to induce the opening of the corresponding hand. All patients achieved high control accuracies and exhibited improvements in motor function. In a later study, Cho et al. [23] present the results of two patients who performed the BCI training with first-person avatar feedback. After the study, both patients reported improvements in motor functions and both have improved their scores on Upper Extremity Fugl-Meyer Assessment scale. Even if the number of patients presented in these two studies is low, they support the idea that this kind of systems may bring additional benefits to the rehabilitation process outcome in stroke patients.

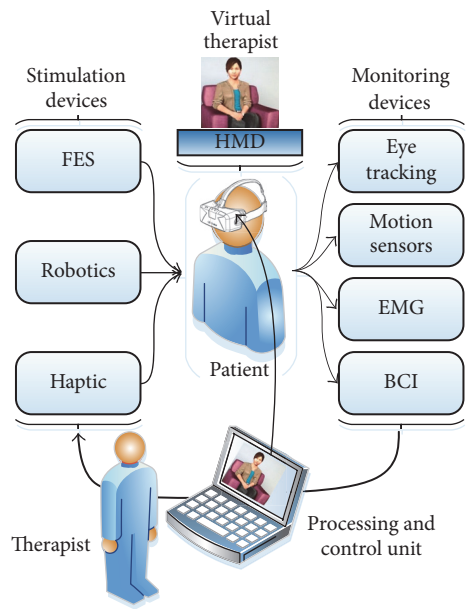


Figure 1: TRAVEE system architecture.

2. General System Architecture

The BCI-FES technique presented in this paper is part of a much more complex system designed for stroke rehabilitation called TRAVEE [24], presented in Figure 1. The stimulation devices, the monitoring devices, the VR headset, and a computer running the software are the main modules of the TRAVEE system. The stimulation devices help the patient to perform the exercises and the monitoring devices are used to determine if the exercises are executed properly, according to the proposed scenarios. Actually, the TRAVEE system must be seen as a software kernel that allows defining a series of rehabilitation exercises using a series of USB connectable devices. This approach is very useful because it offers the patient the options to buy, borrow, or rent the abovementioned devices according to his needs and after connection, the therapist may choose the suitable set of exercises.

The TRAVEE system is based on a new and promising rehabilitation concept which implies the augmented/magnified feedback of the movement of the impaired limb and can be successfully applied especially in the early stages of the rehabilitation therapy in order to close the loop that may trigger the mirror neurons [25]. These mirror neurons intermediate learning, indirectly controlling the brain plasticity and the technique is known as mirror therapy for stroke rehabilitation [26]. Despite the advantages of mirror therapy in comparison with other standard techniques, some disadvantages are obvious: it is difficult to explain to a patient how the mirror helps him: monotony, the patient's condition

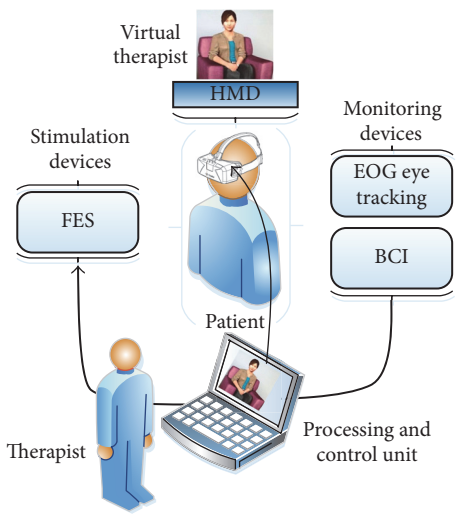


Figure 2: The BCI-FES TRAVEE subsystem.

Figure 3: The hand rehabilitation exercise.

and position, the lack of challenging task, and so on. [27]. By replacing the physical mirror with a VR headset the patient has the same visual feedback that is needed to close the loop that triggers the mirror neurons but without disadvantages of the mirror therapy mentioned above. Once the patient is immersed in the virtual world he is no longer a disabled person and this has a good impact on patient's self-esteem. Within the TRAVEE project, encouraging results were obtained for the development of a virtual reality system for poststroke recovery using an inertial movement unit, a glove with sensors, a Myo Armband with electromyography sensors, and an Oculus Rift headset [28]. An alternate implemented system contains a Leap Motion device for patient's limbs movements monitoring, a VR headset, and a haptic module attached to patient's arm also offering better results than standard therapy methods [29].

3. Materials and Methods

For the current study, the BCI-FES TRAVEE subsystem is composed of FES as stimulation device, BCI and an electrooculography (EOG) system as monitoring devices, Oculus Rift as VR headset, and a laptop, Figure 2.

The rehabilitation exercise was focused on flexion and extension of hand and fingers (Figure 3). The patient is seated in a wheelchair or normal chair. The FES electrodes are

Figure 4: Patient executing a rehabilitation exercise.

mounted on extensors muscles of both hands as shown in Figure 3 and the FES software module is started in order to determine the FES parameters (intensity and timings of the current impulse: rising, front, and falling). Then, the EOG electrodes and EEG helmet are mounted and the correct acquisition of the signals is verified. Before attaching the VR headset, the therapist sits in front of the patient explaining what he will see by showing him the following: the virtual therapist will raise the hand like in Figure 3 (the left hand of the therapist is the right hand of the patient); a big arrow will appear on the upper left or right of the screen depending on virtual therapist indications and the patient will also hear sounds from the left or the right. After explanations, the VR headset is mounted on (Figure 4), EOG system is calibrated, and the recovery exercise may begin, but not before the real therapist tells the patient that he has the possibility of choosing between two views: front view (the virtual therapist is located in front of the patient) or mirror view (the virtual therapist is located on the left side and a mirror is in front of them, like in a dance room) presented in Figure 5.

For the EOG calibration, a red spot appears for 2.5 seconds on a white background displayed on the VR system in different places, in the following order: center, upper right, center, upper left, center, lower left, center, lower right, and center. The user has to gaze at the spot in each location. The calibration is very important for an accurate calculation of the gaze points (eye tracking) during the tests.

In order to provide VR and FES feedback according to the patient's imagined movement, a set of spatial filters and classifier have to be created [22]. First, we are recording 4 runs of training data. Each run consists of 20 right- and 20 left-MI trials, in a random order. We use the trial time course and signal processing algorithms presented in [22]. Each trial lasts 8 seconds. At second 2 a beep informs the user about the upcoming cue. At second 3, the cue is presented and marks the moment when the user has to start imagining the movement shown by the virtual therapist until the end of the trial. While recording the test data, starting with second 4.25, the user sees the virtual hand indicated by the cue moving, and at the same time, the neurostimulator induces the patient's corresponding hand opening. After the spatial filters and classifier are created, we are recording 2 more runs, where the VR and FES feedback are provided to the patient between seconds 4.25 and 8 of each trial only if the classification result is correct. By comparing every sample of the classification result with the presented cue for each

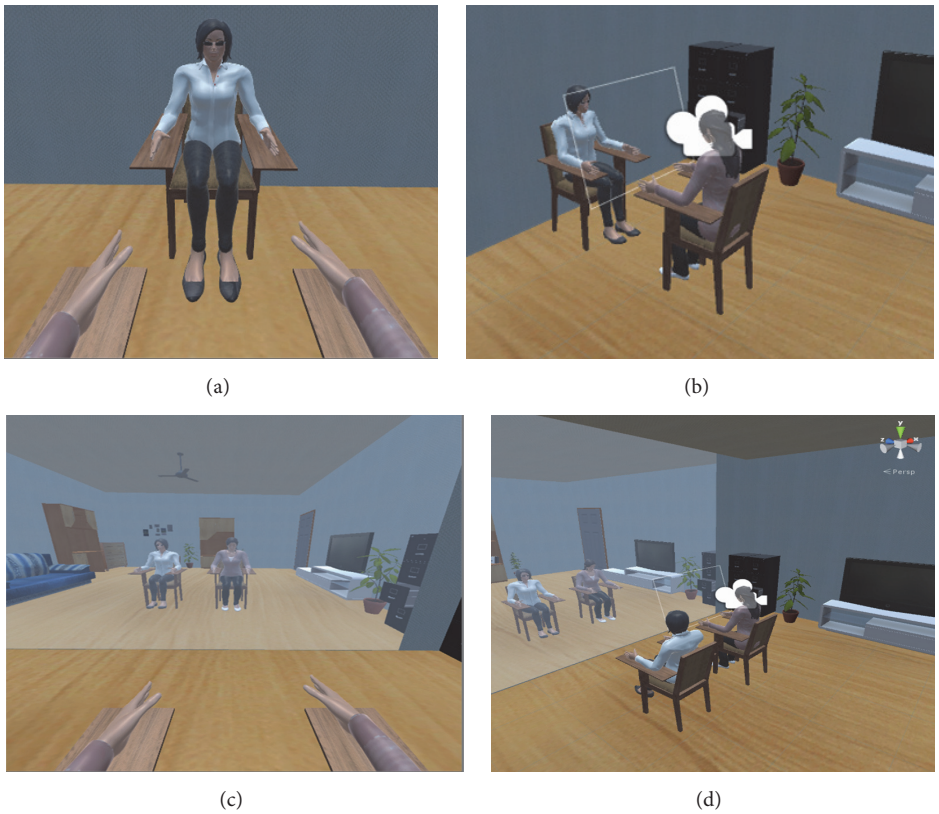


Figure 5: The VR environment in which the patient is immersed: (a) and (c) patient views; (b) and (d) world views; (a) the therapist in front of the patient; (c) the therapist on the left side of the patient with mirror in the front.

trial during the last 2 runs, we are calculating a control error rate course for that session. Except the first session, while recording the 4 train data runs, we are using the set of spatial filters and classifier calculated in the previous session of that patient only if the control error rate for that session was smaller than 20%.

4. EEG and EOG Recording

The BCI-FES subsystem consists of a 16-channel biosignal amplifier (g.USBamp, g.tec medical engineering GmbH) and an 8-channel neurostimulator (MOTIONSTIM8, KRAUTH+ TIMMERMANN GmbH). The EEG signals are collected from 12 positions over the sensorimotor areas according to the 10–20 International System, as seen in Figure 6(a). The last four channels are used in differential mode to record the vertical and horizontal EOG. Figure 6(b) presents the EOG electrodes position of the subject's head. The EEG and EOG data are sampled at 256 Hz and notch-filtered for excluding the 50 Hz noise. The EEG data are bandpass filtered between 8 and 30 Hz and then fed to the processing algorithm that performs spatial filtering with the Common Spatial Patterns (CSP) method [30, 31] and Linear Discriminant Analysis (LDA) classification [22, 32]. The EOG data are filtered with a moving average filter in order to calculate the average of the last 128 samples.

To acquire EOG signals the same EEG device was used but from all the EEG electrodes of the gTec–g.USBamp, 4 of them were used for EOG signals. The eye tracking is necessary because patient needs constant motivation and attention during training/recovering session from a therapist. In fact, after a while, the patient does not pay attention any more, is falling asleep, or is looking at/thinking of something else. By using the electrooculography (EOG) based eye tracking, the system is able to determine if the patient is concentrated and warns the patient if he is not. Figure 7 presents the output of the implemented algorithm for detecting the gaze point of the subject on the image in front of him. Figure 7(a) shows the processed HEOG and VEOG while Figure 7(b) displays the movement of the gaze point based on HEOG and VEOG.

5. Technical and Clinical Testing

The online signal processing and classification of the EEG signals were done by using the Common Spatial Patterns 2 class BCI Simulink model provided by g.tec medical engineering GmbH and the offline analysis of the data was done using g.BSanalyze software provided by the same company. For the EOG processing we developed a Simulink block containing an algorithm that processes the EOG signals and outputs the x-y gaze normalized coordinates with respect to the center point of the image displayed on the VR system. The whole

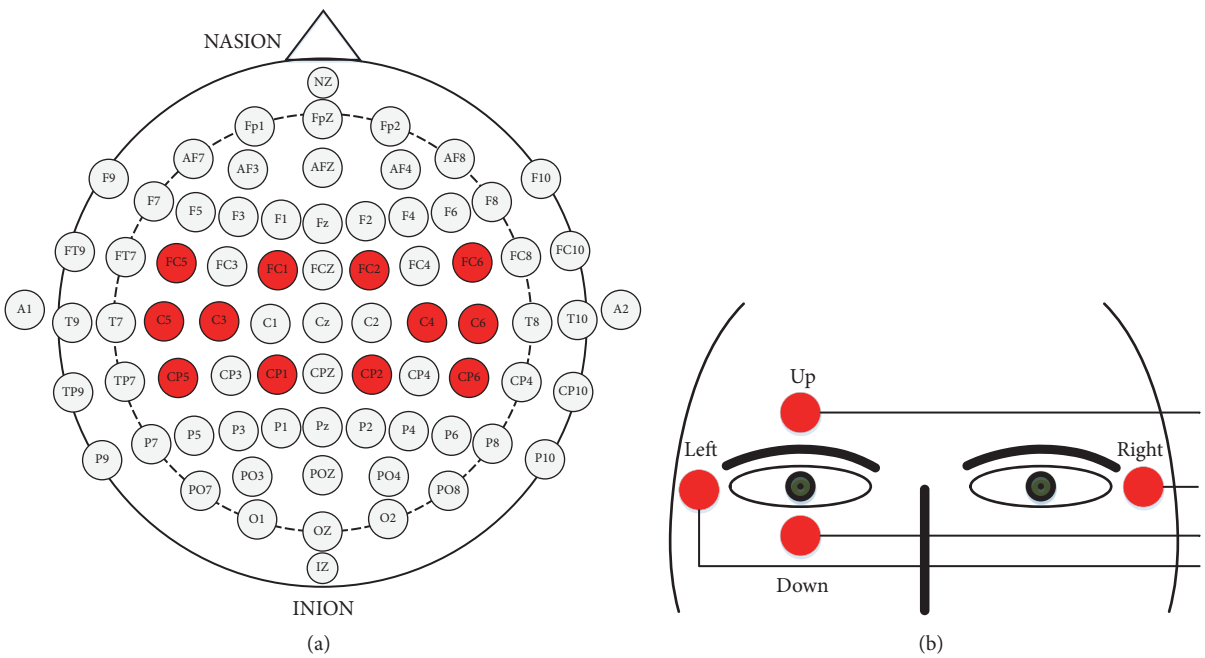


Figure 6: (a) EEG electrodes positions according to the 10–20 International System; (b) EOG electrodes displacement.

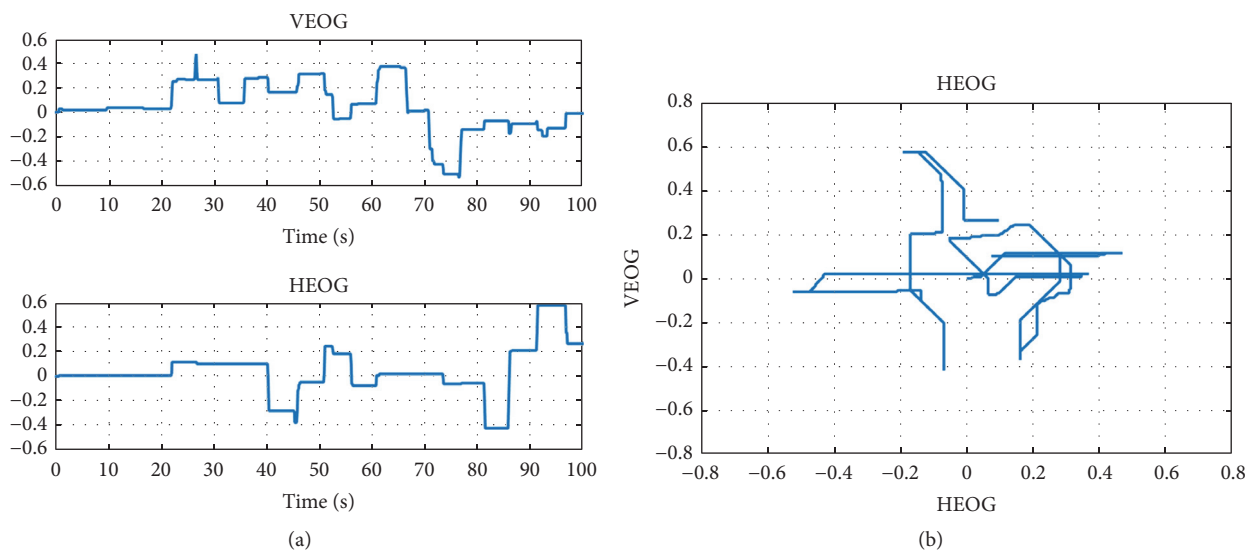


Figure 7: (a) HEOG and VEOG recorded for 100 seconds; (b) the gaze position on the image during 100 seconds of recording.

system was first tested on 3 healthy people and then some fine tunings were done based on their suggestions in order to get high accuracy and a good repeatability coefficient. All three-healthy people achieved low control error rates, comparable to the ones presented by Ortner and colleagues in [33].

Before starting the tests on patients within clinical environment, this study was approved by the institutional review board of the National Institute of Rehabilitation, Physical Medicine and Balneoclimatology from Bucharest, Romania, and each patient signed informed consent and an authorization for videos and photographs release before starting the study. The general clinical profile of the patients included in the study was afebrile, aware, temporospatial oriented, and cardiorespiratory balanced, without digestive or reno-urinary complains, with poststroke central neuromotor syndrome. From the whole patients, one-third was women and two-thirds were men, with ages between 52 and 79 years old. The inclusion criteria was stable neurological status; stable consciousness state; significant and persistent neuromotor deficit; disability for at least two of the following: mobility, self-help capacity, communication, sphincter control, deglutition; sufficient cognitive functions to allow learning; communication ability; sufficient physical exercise tolerance.

The clinical tests are in progress and until this moment the proposed system was tested on 7 patients. Each of them performed three training sessions, and all of them were able

Table 1: Mean and minimal control error rate values for seven patients.

Subject	Session	Mean error [%]	Minimal error [%]
S1	1	20.62	5.48
	2	20.62	7.11
	3	26.48	19.70
S2	1	23.96	11.97
	2	24.60	14.10
	3	28.83	21.00
S3	1	33.56	22.78
	2	37.00	21.35
	3	35.58	29.51
S4	1	32.58	24.77
	2	31.54	24.61
	3	37.21	26.22
S5	1	18.50	7.36
	2	19.72	10.72
	3	20.80	9.45
S6	1	19.20	6.37
	2	19.25	7.68
	3	19.58	1.95
S7	1	28.19	15.00
	2	25.53	13.56
	3	21.91	5.13
Mean values		*25.96*	*14.56*

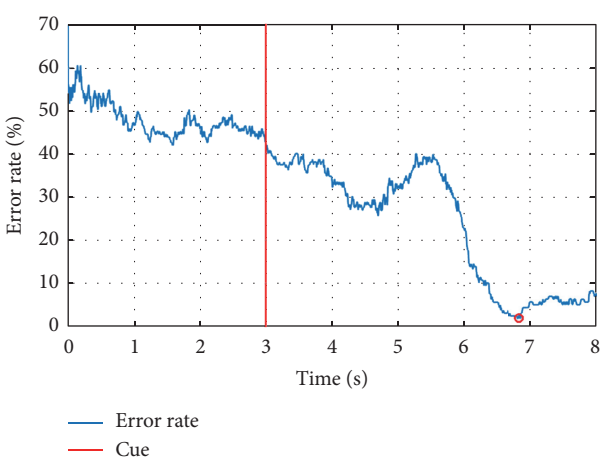


Figure 8: The error rate in time for subject S6, session 3.

to achieve a low control error rate over the whole system. Table 1 presents the mean and minimal control error rate achieved by each patient. The mean error rate is calculated as the mean of the errors for each time point between seconds 4.25 and 8 of the last 2 runs. Figure 8 presents the error rate in time for subject S6, session 8, when he achieved the lowest control error rate, indicated by the red circle at second 6.8.

Except for subjects S3 and S4, all patients exhibited control error rates lower than 20% in at least one session. At this time of the study, it is premature to make evaluations of the rehabilitation outcome of the patients, but, based on their feedback after each session, the VR system makes them remain focused on the task that they have to perform, and they see everything like an interactive game. The fact that they are cognitively involved in this task, unlike having a passive or bored attitude, obviously brings additional benefits to rehabilitation process outcome.

At the beginning, it was difficult for the patients to understand how to concentrate on imagining the movement of their impaired limb as part of the rehabilitation exercise. For those with a low-level education, it was unclear how such a concentration effort regarding their limb movement will help them. This was observed especially when the system was used only with BCI module without VR. The indications on what they had to do were very poor in information (just a simple sound and an arrow to indicate left or right). Also, the activity around the patient disturbed him very easily from imagining the movement. The patients needed around 5 training sessions in order to learn how to imagine the movements and to obtain a good neurofeedback. By adding VR, the number of training sessions was decreased to one or (very rarely) two.

Analyzing the questionnaires, it was concluded that the average user satisfaction was around 3, the answers being highly influenced by the patients' understanding of the rehabilitation therapy because most of them expected to recover themselves based on the therapist's activity and not to be consciously involved in the rehabilitation process. That

depends also on the education degree. However, the overall patients' impression was that they felt and saw an encouraging improvement in recovering after using the proposed system.

For the next months, we plan to organize two groups of patients: a test group and a control group. The test group will perform up to 25 sessions of training with the system, while the control group will perform only classical rehabilitation therapy. When finishing the study, the results will be compared between groups and a statistical analysis will be performed on the results to see if the test group function improvements are statistically and significantly higher than the ones of the control group.

6. Conclusions

In this paper, a BCI-FES system for stroke rehabilitation is presented. Besides stimulation device, the BCI and EOG systems supervise how exercises are performed and the patient's commitment and Oculus Rift headset facilitates the patient's immersion in VR. By using this system, the patient is not distracted by the real environment or by events around him. He is just immersed in VR where the virtual therapist tells and shows him how to perform every exercise and a red big arrow is shown every time. The patient is focused most of the time, but if he loses his concentration the eye tracking system detects this and gives a warning.

The technical performances were validated by testing the system on healthy persons with good knowledge in assistive technologies. The healthy people achieved low control error rates, comparable to the ones reported in the literature.

The clinical tests are in progress, but the preliminary ones are very encouraging regarding fast accommodation and satisfaction of each patient. This approach of combining VR and BCI and FES facilities can effectively speed up the rehabilitation period and increase the users' optimism and the desire to exercise and recover lost skills. By involving the brain via BCI and VR the system proved to be more effective than the standard techniques.

The clinical tests last for several months for a significant number of subjects but once these will be completed the Likert questionnaires and technical files of all subjects will be analyzed.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Romanian National Authority for Scientific Research (UEFISCDI), Project 1/2014 Virtual Therapist with Augmented Feedback for Neuromotor Recovery (TRAVEE).

References

[1] Stroke Statistics, The Internet Stroke Center, http://www.strokecenter.org/patients/about-stroke/stroke-statistics/, last visit September 2017.

[2] Recovering After a Stroke: A Patient and Family Guide, http://www.strokecenter.org/wp-content/uploads/2011/08/Recovering-Aftera-Stroke.pdf, last visit September 2017.

[3] Johns Hopkins Institute - Strock Centers, http://www.hopkinsmedicine.org/neurology_neurosurgery/centers_clinics/cerebrovascular/stroke/.

[4] ENIGMA-Stroke Recovery, http://enigma.ini.usc.edu/ongoing/enigma-stroke-recovery/, last visit September 2017.

[5] StrokeBack Project, http://www.strokeback.eu/project.html, last visit September 2017.

[6] Stroke recovery gets a boost from Kinect last visit September, http://www.microsoft.com/en-us/research/blog/stroke-recovery-gets-a-boost-from-kinect/.

[7] T. Bradshaw, "Virtual Reality gets its mass-market headset on, Financial Times," https://www.ft.com/content/f8087e6e-8c66-11e6-8aa5-f79f5696c731 last visit.

[8] A. Lele, "Virtual reality and its military utility," *Journal of Ambient Intelligence and Humanized Computing*, vol. 4, no. 1, pp. 17–26, 2013.

[9] K. Laver, S. George, S. Thomas, J. E. Deutsch, and M. Crotty, "Virtual reality for stroke rehabilitation: an abridged version of a Cochrane review," *European Journal of Physical and Rehabilitation Medicine*, vol. 51, no. 4, pp. 497–506, 2015.

[10] B. H. Dobkin and A. Dorsch, "New evidence for therapies in stroke rehabilitation.," *Current Atherosclerosis Reports*, vol. 15, no. 6, p. 331, 2013.

[11] W.-P. Teo, M. Muthalib, S. Yamin et al., "Does a combination of virtual reality, neuromodulation and neuroimaging provide a comprehensive platform for neurorehabilitation?—A narrative review of the literature," *Frontiers in Human Neuroscience*, vol. 10, article no. 284, 2016.

[12] Benefits of Virtual Reality for Stroke Rehabilitation, http://www.saebo.com/benefits-virtual-reality-stroke-rehabilitation/, last visit September 2017.

[13] L. Panjwani, Virtual Reality Therapy Designed to Help Stroke Patients Recover last visit September, http://www.rdmag.com/article/2017/08/virtual-reality-therapy-designed-help-stroke-patients-recover.

[14] Virtual Reality in Stroke Rehabilitation at NYDNR, https://nydnrehab.com/treatment-methods/neurorehab/virtual-reality-in-stroke-rehabilitation/.

[15] Stroke Outcomes Research Canada, sorcan.ca/current-projects/,.

[16] N. Sabathiel, D. C. Irimia, B. Z. Allison, C. Guger, and G. Edlinger, "Paired associative stimulation with brain-computer interfaces: A new paradigm for stroke rehabilitation," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 9743, pp. 261–272, 2016.

[17] K. K. Ang, C. Guan, K. S. G. Chua et al., "A large clinical study on the ability of stroke patients to use an EEG-based motor imagery brain-computer interface," *Clinical EEG and Neuroscience*, vol. 42, no. 4, pp. 253–258, 2011.

[18] F. Pichiorri, F. De Vico Fallani, F. Cincotti et al., "Sensorimotor rhythm-based brain-computer interface training: The impact on motor cortical responsiveness," *Journal of Neural Engineering*, vol. 8, no. 2, Article ID 025020, 2011.

[19] R. Ortner, D. C. Irimia, J. Scharinger, and C. Guger, "A motor imagery based brain-computer interface for stroke rehabilitation," *Stud Health Technol Inform*, vol. 181, pp. 319–323, 2012.

[20] S. R. Soekadar, N. Birbaumer, M. W. Slutzky, and L. G. Cohen, "Brain-machine interfaces in neurorehabilitation of stroke," *Neurobiology of Disease*, vol. 83, pp. 172–179, 2015.

[21] A. Remsik, B. Young, R. Vermilyea et al., "A review of the progression and future implications of brain-computer interface therapies for restoration of distal upper extremity motor function after stroke," *Expert Review of Medical Devices*, vol. 13, no. 5, pp. 445–454, 2016.

[22] D. C. Irimia, M. S. Poboroniuc, R. Ortner, B. Z. Allison, and C. Guger, "Preliminary results of testing a BCI-controlled FES system for post-stroke rehabilitation," in *Proceedings of the 7th Graz Brain-Computer Interface Conference 2017, September 18th – 22nd*, Graz, Austria, 2017.

[23] W. Cho, A. Heilinger, R. Xu et al., "Hemiparetic Stroke Rehabilitation Using Avatar and Electrical Stimulation Based on Non-invasive Brain Computer Interface," *International Journal of Physical Medicine & Rehabilitation*, vol. 05, no. 04, 2017.

[24] TRAVEE, Virtual Therapist with Augmented Feedback for Neuromotor Recovery, http://travee.upb.ro/, last visit september 2017.

[25] D. Cinteza, "Modern Concepts of Recovery and Rehabilitation - CNS Affections (Mirror System)," *Balneo-Research Journal*, vol. 3, 2012.

[26] M. E. Michielsen, R. W. Selles, J. N. van der Geest et al., "Motor recovery and cortical reorganization after mirror therapy in chronic stroke patients: a phase II randomized controlled trial," *Neurorehabilitation and Neural Repair*, vol. 25, no. 3, pp. 223–233, 2011.

[27] T. Muzaffar, R. K. Wadhwa, B. Diganta, N. Laisram, and SY. Kothari, "Evaluation of Mirror Therapy for Upper Limb Rehabilitation in Stroke," in *Vol 24(3): 63-9*, vol. 24, p. 63, September, IJPMR, 2013.

[28] R. G. Lupu, F. Ungureanu, and A. Stan, "A virtual reality system for post stroke recovery," in *Proceedings of the 20th International Conference on System Theory, Control and Computing, ICSTCC 2016*, pp. 300–305, Romania, October 2016.

[29] R. G. Lupu, N. Botezatu, F. Ungureanu, D. Ignat, and A. Moldoveanu, "Virtual reality based stroke recovery for upper limbs using leap motion," in *Proceedings of the 20th International Conference on System Theory, Control and Computing, ICSTCC 2016*, pp. 295–299, Romania, October 2016.

[30] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clinical Neurophysiology*, vol. 110, no. 5, pp. 787–798, 1999.

[31] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.

[32] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *NeuroImage*, vol. 56, no. 2, pp. 387–399, 2011.

[33] R. Ortner, J. Scharinger, A. Lechner, and C. Guger, "How many people can control a motor imagery based BCI using common spatial patterns?" in *Proceedings of the 7th International IEEE/EMBS Conference on Neural Engineering (NER '15)*, pp. 202–205, April 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 3918284, 17 pages
https://doi.org/10.1155/2018/3918284

Review Article

Current Use and Future Perspectives of Spatial Audio Technologies in Electronic Travel Aids

Simone Spagnol,[1,2] György Wersényi,[3] Michał Bujacz,[4] Oana Bălan,[5] Marcelo Herrera Martínez,[2,6] Alin Moldoveanu,[5] and Runar Unnthorsson[2]

[1]*Department of Information Engineering, University of Padova, Via Gradenigo 6B, 35131 Padova, Italy*
[2]*Department of Industrial Engineering, Mechanical Engineering and Computer Science, University of Iceland, Reykjavík, Iceland*
[3]*Department of Telecommunications, Széchenyi István University, Györ, Hungary*
[4]*Institute of Electronics, Technical University of Lodz, Lodz, Poland*
[5]*Faculty of Automatic Control and Computers, Politehnica University of Bucharest, Bucharest, Romania*
[6]*Faculty of Engineering, University of San Buenaventura, Carrera 8H #172-20, Bogotá, Colombia*

Correspondence should be addressed to Simone Spagnol; spagnols@hi.is

Received 23 August 2017; Revised 23 January 2018; Accepted 12 February 2018; Published 21 March 2018

Academic Editor: Tao Gu

Electronic travel aids (ETAs) have been in focus since technology allowed designing relatively small, light, and mobile devices for assisting the visually impaired. Since visually impaired persons rely on spatial audio cues as their primary sense of orientation, providing an accurate virtual auditory representation of the environment is essential. This paper gives an overview of the current state of spatial audio technologies that can be incorporated in ETAs, with a focus on user requirements. Most currently available ETAs either fail to address user requirements or underestimate the potential of spatial sound itself, which may explain, among other reasons, why no single ETA has gained a widespread acceptance in the blind community. We believe there is ample space for applying the technologies presented in this paper, with the aim of progressively bridging the gap between accessibility and accuracy of spatial audio in ETAs.

1. Introduction

Spatial audio rendering techniques have various application areas ranging from personal entertainment, through teleconferencing systems, to real-time aviation environments [1]. They are also used in health care, for instance, in motor rehabilitation systems [2], electronic travel aids (ETAs, i.e., devices which aid in independent mobility through obstacle detection or help in orientation and navigation) [3], and other assistive technologies for visually impaired persons [4].

In the case of ETAs, the hardware has to be portable, lightweight, and user-friendly, allow for real-time operation, and be able to support long-term operation. All these issues put designers and developers to a challenge where state-of-the-art technology literally comes at hand in the form of high-tech mobile devices, smartphones, and so on. Furthermore, if ETAs are designed for the visually impaired (The term Electronic Travel Aid was born and is almost exclusively used to describe systems developed to help visually impaired persons with navigating their surroundings safely and efficiently. Nevertheless, visually impaired persons are not strictly the only group who might benefit from ETAs: for instance, non-visual interaction focused towards navigation is of interest to firefighters operating in smoke-filled buildings [5].), even more aspects have to be considered. Beyond the aforementioned, the devices should have a special user interface as well as alternative input and output solutions, where feedback in the form of sound can enhance the functionality of the device. Most of the developments of ETAs for the visually impaired aim at safety during navigation, such as avoiding obstacles, recognizing objects, and extending the auditory information by spatial cues [6, 7]. Since visually impaired persons rely on spatial audio cues as their primary sense of orientation [8], providing them with an accurate virtual auditory representation of the environment is essential.

ETAs evolved considerably over the past years, and a variety of virtual auditory displays [9] were proposed, using different spatial sound techniques and sonification approaches, as well as basic auditory icons, earcons, and speech [10]. Available ETAs for the visually impaired provide various information that ranges from simple obstacle detection with a single range-finding sensor, to more advanced feedback employing data generated from visual representations of the scenes, acquired through camera technologies. The auditory outputs of such systems range from simple binary alerts indicating the presence of an obstacle in the range of a sensor, to complex spatial sound patterns aiming at sensory substitution and carrying almost as much information as a graphical image [7, 11].

A division can also be made between *local mobility aids* (environmental imagers or obstacle detectors, with visual or ranging sensors) that present only the nearest surroundings to the blind traveler and *navigation aids* (usually GPS- or beacon-based) that provide information on path waypoints [12] or geographical points of interest [13]. While the latter group focuses on directions towards the next waypoint, meaning that a limited spatial sound rendering could be used (e.g., just presenting sounds in the horizontal plane) [14], the former group primarily provides information on obstacles (or the lack of them) and near scene layouts (e.g., walls and shorelines), supporting an accurate spatial representation of the scene [6].

Nevertheless, most of these systems are still in their infancy and at a prototype stage. Moreover, no single electronic assistive device has gained a widespread acceptance in the blind community, for different reasons: limited functionalities, ergonomics, small scientific/technological value, limited end-user involvement, high cost, and potential lack of commercial/corporate interest in pushing high-quality electronic travel aids [3].

While many excellent recent reviews on ETA solutions are available (see, e.g., [3, 4, 6, 7]), to our knowledge none of these works critically discusses or analyzes in depth the important aspect of spatial audio delivery. This paper gives an overview about existing solutions for delivering spatial sound, focusing on wearable technologies suitable for use in electronic travel aids for the visually impaired. The analysis reported in this paper indicates a significant potential to achieve accurate spatial sound rendering through state-of-the-art audio playback devices suitable for visually impaired persons and advances in customization of virtual auditory displays. This review was carried out within the European Horizon 2020 project named Sound of Vision (http://www.soundofvision.net). Sound of Vision focuses on creating an ETA for the blind that translates 3D environment models, acquired in real-time, into their corresponding real-time auditory and haptic representations [15].

The remainder of the paper is organized as follows. Section 2 reviews the basics of 3D sound localization, with a final focus on blind localization. Section 3 introduces the available state-of-the-art software solutions for customized binaural sound rendering, while Section 4 presents the available state-of-the-art hardware solutions suitable for the visually impaired. Finally, in Section 5 we discuss current uses and future perspectives of spatial audio in ETAs.

2. Basics of 3D Sound Localization

Localizing a sound source means determining the location of the sound's point of origin in the three-dimensional sound space [16]. Location is defined according to a head-related coordinate system, for instance, the interaural polar system. In the interaural polar coordinate system the origin coincides with the interaural midpoint and the elevation angle ϕ goes from -180° to 180° with negative values below the horizontal plane and positive values above, while the azimuth angle θ ranges from -90° at the left ear to 90° at the right ear. The third dimension, distance r, is the Euclidean distance between the sound source and the origin. In the following we will refer to the three planes that divide the head into halves as the *horizontal* plane (upper/lower halves), the *median* plane (left/right halves), and the *frontal* plane (front/back halves).

Spatial cues for sound localization can be categorized according to polar coordinates. As a matter of fact, each coordinate is thought to have one or more dominant cues in a certain frequency range associated with a specific body component, in particular the following:

(i) Azimuth and distance cues at all frequencies are associated with the head.

(ii) Elevation cues at high frequencies are associated with the pinnae.

(iii) Elevation cues at low frequencies are associated with torso and shoulders.

Based on well-known concepts and results, the most relevant cues for sound localization are now discussed [17].

2.1. Azimuth Cues. At the beginning of the twentieth century, Lord Rayleigh studied the means through which a listener is able to discriminate at a first level the horizontal direction of an incoming sound wave. Following his Duplex Theory of Localization [18], azimuth cues can be reduced to two basic quantities thanks to the active role of the head in the differentiation of incoming sound waves, that is, the following:

(i) *Interaural Time Difference* (ITD), defined as the temporal delay between sound waves at the two ears

(ii) *Interaural Level Difference* (ILD), defined as the ratio between the instantaneous amplitudes of the same two sounds.

ITD is known to be frequency-independent below 500 Hz and above 3 kHz, with an approximate ratio of low-frequency ITD by high-frequency ITD of 3/2, and slightly variable at middle range frequencies [19]. Conversely, frequency-dependent shadowing and diffraction effects introduced by the human head cause ILD to greatly depend on frequency.

Consider a low-frequency sinusoidal signal (up to 1 kHz approximately). Since its wavelength is greater than the head dimensions, ITD is no more than a phase lag $\Delta\phi < 2\pi$ between the signals arriving at the ears and therefore a reliable cue for horizontal perception in the low-frequency range [16]. Conversely, the considerable shielding effect of the human

head on high-frequency waves (above 1 kHz) makes ILD the most relevant cue in such spectral range.

Still, the information provided by ITD and ILD can be ambiguous. If one assumes a spherical geometry of the human head, a sound source located in front of the listener at azimuth θ and a second one located at the rear, at azimuth $180 - \theta$, provide in theory identical ITD and ILD values. In practice, ITD and ILD will not be identical at these two azimuth angles because the human head is clearly not spherical, and all subjects exhibit slight asymmetries with respect to the median plane. Nonetheless their values will be very similar, and *front-back confusion* is in fact often observed experimentally [20]: listeners erroneously locate sources at the rear instead of the front (or less frequently, vice versa).

2.2. Elevation Cues. Directional hearing in the median vertical plane is known to have lower resolution compared with that in the horizontal plane [21]. For the record, the smallest change of position of a sound source producing a just-noticeable change of position of the auditory event (known as "localization blur") along the median plane was found to be never less than 4°, reaching a much larger threshold (≈17°) for unfamiliar speech sounds, as opposed to a localization blur of approximately 1°-2° in the frontal part of the horizontal plane for a vast class of sounds [16]. Such a poor resolution is due to

(i) the need of high-frequency content (above 4-5 kHz) for accurate vertical localization [22, 23];

(ii) mild interaural differences between the signals arriving at the left and right ear for sources in the median plane.

If a source is located outside the horizontal plane, ITD- and ILD-based localization becomes problematic. As a matter of fact, sound sources located at all possible points of a conic surface pointing towards the ear of a spherical head produce the same ITD and ILD values. These surfaces, which generalize the aforementioned concept of front-back confusion for elevation angles, are known as *cones of confusion* and represent a potential difficulty for accurate perception of sound direction.

Nonetheless, it is undisputed that vertical localization ability is brought by the presence of the pinnae [24]. Even though localization in any plane involves pinna cavities of both ears [25], determination of the perceived vertical angle of a sound source in the median plane is essentially a monaural process [26]. The external ear plays an important role by introducing peaks and notches in the high-frequency spectrum of the incoming sound, whose center frequency, amplitude, and bandwidth greatly depend on the elevation angle of the sound source [27, 28], to a remarkably minor extent on azimuth [29], and are almost independent of distance between source and listener beyond a few centimeters from the ear [30, 31]. Such spectral effects are physically due to reflections on pinna edges as well as resonances and diffraction inside pinna cavities [26, 29, 32].

In general, both pinna peaks and notches are thought to play an important function in vertical localization of a sound source [33, 34]. Contrary to notches, peaks alone are not sufficient vertical localization cues [35]; however, the addition of spectral peaks supports the improvement of localization performance at upper directions with respect to notches alone [36]. It is also generally considered that a sound source has to contain substantial energy in the high-frequency range for accurate judgement of elevation, because wavelengths significantly longer than the size of the pinna are not affected. Since wavelength λ and frequency f are related as $\lambda = c/f$ (Here c is the speed of sound, typically $c = 343.2$ m/s in dry air at 20°C.), we could roughly state that pinnae have relatively little effect below $f = 3$ kHz, corresponding to an acoustic wavelength of $\lambda \approx 11$ cm.

While the role of the pinna in vertical localization has been extensively studied, the role of torso and shoulders is less understood. Their effects are relatively weak if compared to those due to the head and pinnae, and experiments to establish the perceptual importance of the relative cues have produced mixed results in general [23, 37, 38]. Shoulders disturb incident sound waves at frequencies lower than those affected by the pinna by providing a major additional reflection, whose delay is proportional to the distance from the ear to the shoulder when the sound source is directly above the listener. Complementarily, the torso introduces a shadowing effect for sound waves coming from below. Torso and shoulders are also commonly seen to perturb low-frequency ITD, even though it is questionable whether they may help in resolving localization ambiguities on a cone of confusion [39].

However, as Algazi et al. remarked [38], when a signal is low-passed below 3 kHz, elevation judgement is very poor in the median plane if compared to a broadband source but proportionally improves as the source is progressively moved away from the median plane, where performance is more accurate in the back than in the front. This result suggests the existence of low-frequency cues for elevation that although being overall weak is significant away from the median plane.

2.3. Distance and Dynamic Cues. Distance estimation of a sound source (see [40] for a comprehensive review on the topic) is even more troublesome than elevation perception. At a first level, when no other cue is available, sound intensity is the first variable that is taken into account: the weaker the intensity is, the farther the source should be perceived. Under anechoic conditions, sound intensity reduction with increasing distance can be predicted through the inverse square law: intensity of an omnidirectional sound source will decay by approximately 6 dB for each doubling distance [41]. Still, a distant blast and a whisper at few centimeters from the ear could produce the same sound pressure level at the eardrum. Having a certain familiarity with the involved sound is thus a second fundamental requirement [42].

However, the apparent distance of a sound source is systematically underestimated in an anechoic environment [43]. On the other hand, if the environment is reverberant, additional information can be given by the direct to reflected energy ratio, or DRR, which functions as a stronger cue for distance than intensity: a sensation of changing distance occurs if the overall intensity is constant but the DRR is altered [41]. Furthermore, distance-dependent spectral effects also

have a role in everyday environments: higher frequencies are increasingly attenuated with distance due to air absorption effects.

Literature on source direction perception is generally based on a fundamental assumption; that is, the sound source is sufficiently far from the listener. In particular, previously discussed azimuth and elevation cues are distance-independent when the source is in the so-called *far-field* (approximately more than 1.5 m from the center of the head) where sound waves reaching the listener can be assumed to be planar. On the other hand, when the source is in the *near field* some of the previously discussed cues exhibit a clear dependence on distance. By gradually approaching the sound source to the listener's head in the near field, it was observed that low-frequency gain is emphasized; ITD slightly increases; and ILD dramatically increases across the whole spectrum for lateral sources [20, 30, 44]. The following conclusions were drawn:

(i) Elevation-dependent features are not correlated to distance-dependent features.

(ii) ITD is roughly independent of distance even when the source is close.

(iii) Low-frequency ILDs are the dominant auditory distance cues in the near field.

It should be then clear that ILD-related information needs to be considered in the near field, where dependence on distance cannot be approximated by a simple inverse square law.

Finally, it has to be remarked that, switching from a static to a dynamic environment where the source and/or the listener move with respect to each other, both source direction and distance perception improve. The tendency to point towards the sound source in order to minimize interaural differences, even without visual aid, is commonly seen and aids in disambiguating front/back confusion [45]. Active motion helps especially in azimuth estimation and to a lesser extent in elevation estimation [46]. Furthermore, thanks to the *motion parallax* effect, slight translations of the listener's head on the horizontal plane can help discriminate source distance [47, 48]: if the source is near, its angular direction will drastically change after the translation (reflecting itself onto interaural differences), while for a distant source this will not happen.

2.4. Sound Source Externalization. Real sound sources are typically *externalized*, that is, perceived to be located outside our own head. However, when virtual 3D sound sources are presented through headphones (see next section), in-the-head localization may typically occur and have a major impact on localization ability. Alternatively, listeners may perceive the direction of the sound source and be able to make accurate localization judgements yet accompanied with perception of the source being way closer to the head than otherwise intended (e.g., on the surface of the skull [49]). However, when relevant constraints are taken into account, such as the use of individually measured head-related transfer functions as explained in Section 3, virtual sound sources can be externalized almost as efficiently as real sound sources [50, 51]. Externalization is, along with other attributes such as coloration, immersion, and realism, one of the key perceptual attributes that go beyond the basic issue of localization recently proposed for the evaluation of virtually rendered sound sources [52].

In-the-head localization is mainly introduced by the loss of accuracy in interaural level differences and spectral profiles in virtually rendered sound sources [49]. Another extremely important factor is given by the interaural and spectral changes triggered by natural head movements in real-life situations: correctly tracked head movements can indeed substantially enhance externalization in virtual sonic environments, especially for sources close to the median plane (hardest to externalize statically in anechoic conditions, due to minimal interaural differences [53]), and even relatively small movements of a few degrees can efficiently reduce in-the-head localization [54]. Furthermore, it has been recently showed that externalization can persist once coherent head movement with the virtual auditory space is stopped [55].

Finally, factors related to sound reverberation contribute to a strong sense of externalization, as opposed to dry anechoic sound. The introduction of artificial reverberation [56] through image-source model-based early reflections, wall and air absorption, and late reverberation can significantly contribute to sound image externalization in headphone-based 3D audio systems [57], as well as congruence between the real listening room and the virtually recreated reverberating environment [58].

2.5. Auditory Localization by the Visually Impaired. A number of previous studies showed that sound source localization by visually impaired persons can be different from that of sighted persons. It has to be first highlighted that previous investigations on visually impaired subjects indicated neither better auditory sensitivity [59–61] nor lower auditory hearing thresholds [62] compared to normally sighted subjects. On the other hand, visually impaired subjects acquire the ability to use auditory information more efficiently thanks to the plasticity of the central nervous system, as, for instance, in speech discrimination [63], temporal resolution [64], or spatial tuning [65].

Experiments with real sound sources suggest that visually impaired (especially early blind) subjects map the auditory environment with equal or better accuracy than sighted subjects on the horizontal plane [62, 66–68] but are less accurate in detecting elevation [67] and show an overly compressed auditory distance perception beyond the near field [69]. However, unlike sighted subjects, visually impaired subjects can correctly localize sounds monaurally [66, 70], which suggests a trade-off in the localization proficiency between the horizontal and median planes taking place [71]. By comparing behavioral and electrophysiological indices of spatial tuning within the central and peripheral auditory space in congenitally blind and normally sighted but blindfolded adults, it was found that blind participants displayed localization abilities that were superior to those of sighted controls, but only when attending to sounds in peripheral auditory space [72]. Still, it has to be taken into account that early blind subjects have no possibility of learning the mapping between auditory events and visual stimuli [73].

While localizing, adapting to the coloration of the signals is a relevant component for both sighted and blind subjects. Improved obstacle sense of the blind is also mainly due to enhanced sensitivity to echo cues [74], which allows so-called echolocation [75, 76]. Thanks to this obstacle sensing ability, which can be improved by training, distance perception in blind subjects may be enhanced [68, 76–78]. In addition, some blind subjects are able to determine size, shape, or even texture of obstacles based on auditory cues [70, 77, 79, 80].

Switching to virtual auditory displays, that is, the focus of this paper, a detailed comparative evaluation of blind and sighted subjects [81] confirmed some of the previously discussed results in the literature on localization with real sound sources. Better performance in localizing static frontal sources was obtained in the blind group due to a decreased number of front-back reversals. In the case of moving sources, blind subjects were more accurate in determining movements around the head in the horizontal plane. Sighted participants, however, performed better during listening to ascending movements in the median plane and in identifying sound sources in the back. In-the-head localization rates and the ability to detect descending movements were almost identical for the two groups. In a further experiment [82] error rates of about 6 to 14 degrees horizontally and 9 to 24 degrees vertically were measured for a pool of blind subjects. Improvements in localization by blind persons were observed mainly in the horizontal plane and in case of a broadband stimulus.

Finally, although visual information corresponding to auditory information significantly aids localization and creation of correct spatial mental mappings, it has to be remarked that visually impaired subjects can benefit from off-site representations in order to gain spatial knowledge of a real environment. For instance, results of recent studies showed that interactive exploration of virtual acoustic spaces [83–85] and audio-tactile maps [86] can provide relevant information for the construction of coherent spatial mental maps of a real environment in blind subjects and that such mental representations preserve topological and metric properties, with performances comparable or even superior to an actual navigation experience.

3. Binaural Technique

The most basic method for simulating sound source direction over loudspeakers is to use panning. This usually refers to amplitude panning using two channels (stereo panning). In this case, only level information is used as a balance between the channels, and the virtual source is shifted towards the louder channel. However, ILD and spectral cues are determined by the actual speaker locations. In traditional stereo setups, where loudspeakers and listener form a triangle, sources can be correctly simulated on the line ideally connecting the two speakers. However, although traditional headphones also use two channels, correct directional information is not maintained due to a different arrangement of the speakers with respect to the listener and by the loss of crosstalk between the channels.

Spatial features of virtual sound sources can be more realistically rendered through headphones by processing an input sound with a pair of filters, each simulating all the linear transformations undergone by the acoustic signal during its path from the sound source to the corresponding listener's eardrum. These filters are known in the literature as head-related transfer functions (HRTFs) [87], formally defined as the frequency-dependent ratio between the sound pressure level (SPL) $\Phi(\theta, \phi, \omega)$ at the eardrum and the free field SPL at the center of the head $\Phi_f(\omega)$ as if the listeners were absent:

$$H(\theta, \phi, \omega) = \frac{\Phi(\theta, \phi, \omega)}{\Phi_f(\omega)}, \tag{1}$$

where (θ, ϕ) indicates the angular position of the source relative to the listener and ω is angular frequency. The HRTF contains all of the information relative to sound transformations caused by the human body, in particular by the head, external ears, torso, and shoulders.

HRTF measurements are typically conducted in large anechoic rooms. Usually, a set of loudspeakers is arranged around the subject, pointing towards him/her and spanning an imaginary spherical surface. The listener is positioned so that the center of the interaural axis coincides with the center of the sphere defined by the loudspeakers and their rotation (or, equivalently, the subject's rotation). A probe microphone is inserted into each ear, either at the entrance or inside the ear canal. The measurement technique consists in recording and storing the signal arriving at the microphones. Consequently, these signals are processed in order to remove the effects of the room and the recording equipment (especially speakers and microphones), leaving only the HRTF [87, 88].

By processing a desired monophonic sound signal with a pair of individual HRTFs, one per channel, and by adequately accounting for headphone-induced spectral coloration (see next Section), authentic 3D sound experiences can take place. Virtual sound sources created with individual HRTFs can be localized almost as accurately as real sources and efficiently externalized [50], provided that head movements can be made and that the sound is sufficiently long [89]. As a matter of fact, localization of short broadband sounds without head movements is less accurate for virtual sources than for real sources, especially in regard to vertical localization accuracy [90], and front/back reversal rates are higher for virtual sources [89].

Unfortunately, the individual HRTF measurement technique requires the use of dedicated research facilities. Furthermore, the process can take up to several hours, depending on the used measurement system and on the desired spatial grid density, being uncomfortable and tedious for subjects. As a consequence, most practical applications use nonindividual (or generic) HRTFs, for instance, measured on *dummy heads*, that is, mannequins constructed from average anthropometric measurements. Several generic HRTF sets are available online. The most popular are based on measurements using the KEMAR mannequin [91] or the Neumann KU-100 dummy head (see the Club Fritz study [92]). Alternatively, an HRTF set can be taken from one of many public databases of individual measurements (see, e.g., [93]); many of these databases were recently unified in a common HRTF format

known as Spatially Oriented Format for Acoustics (SOFA) (https://www.sofaconventions.org/).

On the other hand, while nonindividual HRTFs represent the cheapest means of providing 3D perception in headphone reproduction, especially in the horizontal plane [94, 95], listening to nonindividual spatial sounds is more likely to result in evident sound localization errors such as incorrect perception of source elevation, front-back reversals, and lack of externalization [96] that cannot be fully counterbalanced by additional spectral cues, especially in static conditions [46]. In particular, individual elevation cues cannot be characterized through generic spectral features.

For the above reasons, different alternative approaches towards HRTF-based synthesis were proposed throughout the last decades [37, 97]. These are now reviewed and presented sorted by increasing level of customization.

3.1. HRTF Selection Techniques. HRTF selection techniques typically use specific criteria in order to choose the best HRTF set for a particular user from a database. Seeber and Fastl [98] proposed a procedure according to which one HRTF set is selected based on multiple criteria such as spatial perception, directional impression, and externalization. Zotkin et al. [99] selected the HRTF set that best matched an anthropometric data vector of the pinna. Geronazzo et al. [100] and Iida et al. [101] selected the HRTF set whose extracted pinna notch frequencies were closest to the hypothesized frequencies of the user according to a reflection model and an anthropometric regression model, respectively.

Similarly, selection can be targeted at detecting a subset of HRTFs in a database that fit the majority of a pool of listeners. Such an approach was pursued, for example, by So et al. [102] through cluster analysis and by Katz and Parseihian [103] through subjective ratings. The choice of the personal best HRTF among this reduced set is left to the user. Even different selection approaches were undertaken by Hwang et al. [104] and Shin and Park [105]. They modeled HRIRs on the median plane as linear combinations of basis functions whose weights were then interactively self-tuned by the listeners themselves.

Results of localization tests included in the majority of these works show a general decrease of the average localization error as well as of the front/back reversal and inside-the-head localization rates using selected HRTFs rather than generic HRTFs.

3.2. Analytical Solutions. These methods try to find a mathematical solution for the HRTF, taking into account the size and shape of the head and torso in particular. The most recurring head model in the literature is that of a rigid sphere, where the response related to a fixed observation point on the sphere's surface can be described by means of an analytical transfer function [106]. Brown and Duda [37] proposed a first-order approximation of this transfer function for sources in the far-field as a minimum-phase analog filter. Near-field distance dependence can be accounted for through an additional filter structure [107].

Although the spherical head model provides a satisfactory approximation to the low-frequency magnitude of a measured HRTF [108], it is far less accurate in predicting ITD, which is actually variable around a cone of confusion by as much as 18% of the maximum interaural delay [109]. ITD estimation accuracy can be improved by considering an ellipsoidal head model that can account for the ITD variation and be adapted to individual listeners [110]. It has to be highlighted, however, that ITD estimation from HRTFs is a nontrivial operation, given the large variability of objective and perceptual ITD results produced by different common calculation methods for the same HRTF dataset [111, 112].

A spherical model can also approximate the contribution of the torso to the HRTF. Coaxial superposition of two spheres of different radii, separated by a distance accounting for the neck, results in the snowman model [113]. The far-field behavior of the snowman model was studied in the frontal plane both by direct measurements on two rigid spheres and by computation through multipole reexpansion [114]. A filter model was also derived from the snowman model [113]; its structure distinguishes the two cases where the torso acts as a reflector or as a shadower, switching between the two filter substructures as soon as the source enters or leaves the torso shadow zone, respectively. Additionally, an ellipsoidal model for the torso was studied in combination with the usual spherical head [38]. Such model is able to account for different torso reflection patterns; listening tests confirmed that this approximation and the corresponding measured HRTF gave similar results, showing larger correlations away from the median plane.

A drawback of these techniques is that since they do not consider the contribution of the pinna, the generated HRTFs match measured HRTFs at low frequencies only, lacking spectral features at higher frequencies [115].

3.3. Structural HRTF Models. According to the structural modeling approach, the contributions to the HRTF of the user's head, pinnae, torso, and shoulders, each accounting for some well-defined physical phenomena, are treated separately and modeled with a corresponding filtering element [37]. The global HRTF model is then constructed by combining all the considered effects [116]. Structural modeling opens to an interesting form of content adaptation to the user's anthropometry, since parameters of the rendering blocks can be estimated from physical data, fitted, and finally related to anthropometric measurements.

Structural models typically assume a spherical or ellipsoidal geometry for both the head and torso, as discussed in the previous subsection. Effective customizations of the spherical head radius given the head dimensions were proposed [117, 118], resulting in a close agreement with experimental ITDs and ILDs, respectively. Alternatively, ITD can be synthesized separately using individual morphological data [119]. An ellipsoidal torso can also be easily customized for a specific subject by directly defining control points for its three axes on the subject's torso [114]. Furthermore, a great variety of pinna models is available in the literature, ranging from simple reflection models [120] and geometric models [121] to more complex physical models that treat the pinna either as a configuration of cavities [122] or as a reflecting surface [29]. Structural models of the pinna, simulating its resonant and reflective behaviors in two separate filter blocks, were also proposed [123–125].

Algazi et al. [93] suggested using a number of one-dimensional anthropometric measurements for HRTF fitting through regression methods or other machine learning techniques. This approach was recently pursued in a number of studies [126–129] investigating the correspondence between anthropometric parameters and HRTF shape. When suitable processing is performed on HRTFs, clear relations with anthropometry emerge. For instance, Middlebrooks [130] reported a correlation between pinna size and center frequencies of HRTF peaks and notches and argued that similarly shaped ears that differ in size just by a scale factor produce similarly shaped HRTFs that are scaled in frequency. Further evidence of the correspondence between pinna shape and HRTF peaks [123, 131, 132] and notches [125, 133, 134] is provided in a number of following works. The use of such knowledge leads to the effective parametrization of structural pinna models based on anthropometric parameters, which suggests an improvement in median plane localization with respect to generic HRTFs [135, 136].

3.4. Numerical HRTF Simulations. Numerical methods typically require as input a 3D mesh of the subject, in particular the head and torso, and include approaches such as finite-difference time domain (FDTD) methods [108], the finite element method (FEM) [137], and the boundary element method (BEM) [138].

Recent literature has focused on the BEM. It is known that high-resolution meshes are needed in order to effectively simulate HRTFs with the BEM, especially for the pinna area. Low mesh resolution results indeed in simulated HRTFs that greatly differ from acoustically measured HRTFs at high frequencies, thus destroying elevation cues [139]. However, as the number of mesh elements grows, memory requirements and computational load grow even faster [140]. Recent works introduced the fast multipole method (FMM) and the reciprocity principle (i.e., interchanging sources and receivers) in order to face BEM efficiency issues [140, 141]. Ultimately, localization performances of simulated HRTFs through the BEM were found to be similar to those observed with acoustically measured HRTFs [142], and databases of simulated HRTFs [143] as well as open-source tools for calculating HRTFs through the BEM given a head mesh as input [144] are available online.

On the other hand, image-based 3D modeling, based on the reconstruction of 3D geometry from a set of user pictures, is a fast and cost-effective alternative to obtaining mesh models [145]. Furthermore, the advent of consumer level depth cameras and the availability of huge computational power on consumer computers open new perspectives towards very cheap and yet very accurate calculation of individualized HRTFs.

4. Headphone Technologies

One of the crucial variables for generating HRTF-based binaural audio is the headphone itself. Headphones are of different types (e.g., circumaural, supra-aural, extra-aural, and in-ear) and can have transfer functions that are far from linear. The main issue with classic headphones is that the transfer function between headphone and eardrum heavily varies from person to person and with small displacements of the headphone itself [146, 147]. Such variation is particularly marked in the high-frequency range where important elevation cues generally lie. As a consequence, headphone playback introduces significant localization errors, such as in-the-head localization, front-back confusion, and elevation shift [148].

In order to preserve the relevant localization cues provided by HRTF filtering during headphone listening, various headphone equalization techniques, usually based on a prefiltering with the inverse of the average headphone transfer function, are used [149]. However, previous research suggests that these techniques are little to no effective when nonindividual (even selected) HRTFs are used [149, 150]. On the other hand, several authors support the use of individual headphone compensation in order to preserve localization cues in the high-frequency range [146, 147].

In the case of travel aids for the visually impaired, additional factors need to be considered in the design and choice of the headphone type. Most importantly, ears are essential to provide information about the environment, and visually impaired persons refuse to use headphones during navigation if these either partially or fully cover the ears, therefore blocking environmental noises. The results of a survey of the preferences of visually impaired subjects for a possible personal navigation device [151] showed indeed that the majority of participants rated headphones worn over the ears as the least acceptable output device, compared to other technologies such as bone-conduction and small tube-like headphones, or even a single headphone worn over one ear. Furthermore, those fully blind had much stronger negative feelings about headphones that blocked ambient sounds than those who were partially sighted.

This important consideration shifts our focus to alternative state-of-the-art solutions for spatial audio delivery such as unconventional headphone configurations, bone-conduction headsets, or active transparent headsets.

4.1. Unconventional Headphone Configurations. The problem of ear occlusion can be tackled by decentralizing the point of sound delivery from the entrance of the ear canal to positions around the ear, with one or more transducers per ear. In this case, issues arise regarding the proper direction and distance of each transducer with respect to the ear canal, as well as their types and dimensions. Furthermore, there is a challenge in the spatial rendering technique in that no research results support the application of traditional loudspeaker-based spatial audio techniques (such as Vector Base Amplitude Panning [152] or Ambisonics [153]) to multispeaker headsets and that traditional HRTF measurements do not match with decentralized speaker positions.

The first attempts in delivering spatial audio through multispeaker headphones were performed by König. A decentralized 4-channel arrangement placed on a pair of circumaural earcups for frontal surround sound reproduction was implemented [154] (an alternative small supra-aural configuration was also proposed [155]). Results showed that this speaker arrangement induces individual direction-dependent pinna

cues as they appear in real frontal sound irradiation in the free field for frequencies above 1 kHz [156]. Psychoacoustic effects introduced by the headphone revealed that frontal auditory events are achieved, as well as effective distance perception [154].

The availability of individual pinna cues at the eardrum is imperative for accurate frontal localization [157]. Accordingly, Sunder et al. [158] later proposed the use of a 2-channel frontal projection headphone which customizes nonindividual HRTFs by introducing idiosyncratic pinna cues. Perceptual experiments validated the effectiveness of frontal headphone playback over conventional headphones with reduced front-back confusions and improved frontal localization. It was also observed that the individual spectral cues created by the frontal projection are self-sufficient for front-back discrimination even with the high-frequency pinna cues removed from the nonindividual HRTF. However, additional transducers are needed if virtual sounds behind the head have to be delivered, and timbre differences with respect to the frontal transducers need to be solved.

Greff and Katz [159] extended the above solutions to a multiple transducer array placed around each ear (8 speakers per ear) recreating the pinna-related component of the HRTF. Simulations and subjective evaluations showed that it is possible to excite the correct localization cues provided by the diffraction of the reconstructed wave front on the listener's own pinnae, using transducer driving filters related to a simple spherical head model. Furthermore, different speaker configurations were investigated in a preliminary localization test, the one with transducers placed at grazing incidence all around the pinna showing the best results in terms of vertical localization accuracy and front/back confusion rate.

Recently, Bujacz et al. [160] proposed a custom headphone solution for a prospective ETA with four proximaural speakers positioned above and below the ears, all slightly to the front. Amplitude panning was then used as spatial audio technique to shift the power of the output sound between pairs of speakers, both horizontally and vertically. Results of a preliminary localization test showed a localization accuracy comparable to HRTF-based rendering through high-quality circumaural headphones, both in azimuth and in elevation.

4.2. Bone-Conduction Headsets. The use of a binaural bone-conduction headset (also known as *bonephones*) is an extremely attractive solution for devices intended for the blind as the technology does not significantly interfere with sounds received through the ear canal, allowing for natural perception of environmental sounds. The typical solution is to place vibrational actuators, also referred to as bone-conduction transducers, on each mastoid (the raised portion of the temporal bone located directly behind the ear) or alternatively on the cheek bones just in front of the ears [161]. Pressure waves are sent through the bones in the skull to the cochlea, with some amount of natural sound leakage through air into the ear canals still occurring.

There are some difficulties in using bone conduction for delivering spatial audio. The first is the risk of crosstalk impeding an effective binaural separation: because of the high propagation speed and low attenuation of sound in the human skull, both the ITD and ILD cues are significantly softened. Walker et al. [162] still observed some degree of spatial separation with interaural cues provided through bone conduction and ear canals either free or occluded, especially relative to ILD. Perceived lateralization is even comparable between air conduction and bone conduction with unoccluded ear canals [163]. However, the degradation relative to standard headphones suggests the difficulty to produce large enough interaural differences to simulate sound sources at extreme lateral locations [162].

The second problem is the need to introduce additional transfer functions for correct equalization of HRTF-based spatial audio: the frequency response of the transducer [164] and the transfer function to the bones themselves, referred to as bone-conduction adjustment function (BAF) [165], which takes into account high-frequency attenuation by the skin [166] and differs between individuals, similar to HRTFs. Walker et al. [167, 168] proposed the use of appropriate bone-related transfer functions (BRTFs) in replacement of HRTFs. Stanley [165] derived individual BAFs from equal-loudness judgements on pure tones, showing that individual BAF adjustments to HRTF-based spatial sound delivery were effective in restoring the spectral cues altered by the bone-conduction pathway. This allowed for effective localization in the median plane by reducing up/down reversals with respect to the BAF-uncompensated stimuli. However, there is no way to measure BAFs empirically, and it is unclear whether the use of a generic, average BAF could lead to the same conclusions.

MacDonald et al. [164] reported similar localization results in the horizontal plane between bone conduction and air conduction, using individual HRTFs as the virtual auditory display and headphone frequency response compensation. Lindeman et al. [169, 170] compared localization accuracy between bone conduction with unoccluded ear canals and an array of speakers located around the listener. The results showed that although the best accuracy was achieved with the speaker array in the case of stationary sounds, there was no difference in accuracy between the speaker array and the bone-conduction device for sounds that were moving, and that both devices outperformed standard headphones for moving sounds.

Finally, Barde et al. [171] recently investigated the minimum discernable angle difference in the horizontal plane with nonindividual HRTFs over a bone-conduction headset, resulting in an average value of 10°. Interestingly, almost all participants reported actual sound externalization.

4.3. Active Transparent Headsets. An active headset is able to detect and process environmental sounds through analog circuits or digital signal processing. One of the most important fields of application of active headsets is noise reduction, where the headset uses active noise control [172, 173] to reduce unwanted sound by the addition of an antiphase signal to the output sound. In the case of ETAs, the environmental signal should not be canceled but provided back to the listener (*hear-through* signal) mixed with the virtual auditory display signal in order for the subject to be aware of the surroundings. Binaural hear-through headsets (in-ear headphones with integrated microphones) are typically used

in augmented reality audio (ARA) applications [174], where a combination of real and virtual auditory objects in a real environment is needed [175].

The hear-through signal is a processed version of the environmental sound and should produce similar auditory perception to natural perception with unoccluded ears. Thus, equalization is needed to make the headset acoustically transparent, since it affects the acoustic properties of the outer ear [176]. The most important problem here is poor fit on the head causing leaks and attenuation problems. The fit of the headphone affects isolation and frequency response as well. Using internal microphones inside the headset in addition to the external ones, a controlled adaptive equalization can be realized [177].

The second basic requirement for a hear-through system is that processing of the recorded sound should have minimal latency [175]. As a matter of fact, when the real signal (leaked to the eardrum) is summed up with the hear-through signal, the delayed version can cause audible comb-filtering effects, especially at lower frequencies where leakage is higher. The audibility of comb-filtering effects depends on both the time and amplitude difference between the hear-through signal and the leaked signal [178]. Using digital realizations, which are preferable over analog circuits in the case of an ETA in terms of both cost and size, suitable latencies of less than 1.4 ms, for which the comb-filtering effect was found to be inaudible when the attenuation of the headset is 20 dB or more, can be achieved with a DSP board [179].

Finally, the hear-through signal should preserve localization cues at the ear canal entrance. Since sound transmission from the microphone to the eardrum is independent of direction whether the microphone is inside or at most 6 mm outside the ear canal [180], having binaural microphones just outside the ear canal entrance is sufficient for obtaining the correct listener-dependent spatial information.

5. Spatial Audio in ETAs

From the multitude of ETAs, two main trends in selecting sound cues can be observed, one to provide very limited yet easily interpretable data, typically from a range sensor, and the other to provide an overabundance of auditory data and let the user learn to extract useful information from it (e.g., the vOICe [181]). A third approach, taken for instance by the authors in the Sound of Vision project [15], is to limit the data from a full-scene representation to just the most useful information, for example, by segmenting the environment and identifying the nearest obstacles or detecting special dangerous scene elements such as stairs. Surveys show that individual preferences among the blind can vary greatly, and all three approaches have users that prefer them [182].

In a recent literature review, Bujacz and Strumiłło [6] classified the auditory display solutions implemented in the most widely known ETAs, either commercially available or in various stages of research and development. Of the 22 considered ETAs, 12 use a spatial representation of the environment. However, breaking the list of ETAs down to obstacle detectors (mostly hand-held) and environmental imagers (mostly head-mounted), ETAs that use a spatial representation almost all belong to the second category. Some of them, such as the vOICe [181], Navbelt [183], SVETA [184], and AudioGuider [185], use stereo panning to represent directions, whereas elevation information is either ignored or coded into sound pitch. ETAs (including works not included in the above cited review) that use HRTFs as the spatial rendering method are now summarized. All of the systems presented in the following are laboratory prototypes.

5.1. Available ETAs Using HRTFs. The *EAV (Espacio Acustico Virtual)* system [186] uses stereoscopic cameras to create a low resolution ($16 \times 16 \times 16$) 3D stereopixel map of the environment in front of the user. Each occupied stereopixel becomes a virtual sound source filtered with the user's individual HRTFs, measured in a reverberating environment. The sonification technique employs spatial audio cues (synthesized with HRTFs) and a distance-to-loudness encoding. Sounds were presented through a pair of individually equalized Sennheiser HD-580 circumaural headphones. Classic localization tests with the above virtual auditory display and tests with multiple sources were performed on 6 blind and 6 normally sighted subjects. Subjects were accurate in identifying the objects' position and recognizing shapes and dimensions within the limits imposed by the system's resolution.

The *cross-modal ETA* device [187] is a wearable prototype that consists of low-cost hardware: earphones (no further information provided), sunglasses fitted with two CMOS micro cameras, and a palm-top computer. The system is able to detect the light spot produced by a laser pointer, compute its angular position and depth, and generate a corresponding sound to the position and distance of the pointed surface. The sonification encoding uses directional auditory cues provided through Brown and Duda's structural HRTF model [37], and distance cues through loudness control and reverberation effects. The subjective effectiveness of the sonification technique was evaluated by several volunteers who were asked to use the system and report their opinions. The overall result was satisfactory, with some problems related to the lack of elevation perception. Targets very high and very low were perceived correctly, whereas those laying in the middle were associated with wrong elevations.

The *Personal Guidance System* [12] receives information from a GPS receiver and was evaluated in five different types of configurations involving different types of auditory displays, spatial sound delivery methods (either via classic headphones or through a speaker worn on the shoulder), and tracker locations. No details about the binaural spatialization engine or the headphones used were provided. Fifteen visually impaired subjects traveled a 50 m long pathway with each of the 5 configurations. Results showed that the configuration using binaurally spatialized virtual speech led to the shortest travel times and highest subjective ratings. However, there were many negative comments about the headphones blocking environmental sounds.

The SWAN system [8, 188] aids navigation and guidance through a set of navigation beacons (earcon-like sounds), object-related sounds (provided through spatial auditory icons), location information, and brief prerecorded speech

samples. Sounds are updated in real-time by tracking the subject's orientation and accordingly spatialized through nonindividual HRTFs. Sounds were played either through a pair of Sony MDR-7506 closed-ear headphones or an equalized bone-conduction headset (see [165]). In an experimental procedure, 108 sighted subjects were required to navigate three different maps. Results showed good navigation skills for almost all the participants in both time and path efficiency.

The main idea of the *Virtual Reality Simulator for the visually impaired people* [189] consists in calculating the distance between the user and nearby objects (depth map) and converting it into sound. The depth map is transformed into a spatial auditory map by using 3D sound cues synthesized with individually measured HRTFs from 1003 positions in the frontal field. Sounds were provided through a standard pair of stereophonic headphones (no further information provided). The Virtual Reality Simulator proved to be helpful for visually impaired people in different research experiments performed indoors and outdoors, in virtual and real-life situations. Among the main limitations of the simulator are tracking accuracy and the lack of a real-time HRTF convolver.

The *Real-Time Assistance Prototype* [190], an evolution of the CASBliP prototype [191], encodes objects' position in space based on their distance (inversely proportional to sound frequency), direction (3D binaural sounds synthesized with nonindividual HRTFs), and speed (proportional to pitch variation). Nonindividual HRTFs of a KEMAR mannequin were measured for different spatial points in a 64° azimuth range, a 30° elevation range, and a 15 m distance range. Sounds were provided through a pair of SONY MDR-EX75SL in-ear headphones. Two experiments were performed with four totally blind subjects, one requiring subjects to identify the sound direction and the other one to detect the position of a moving source and to follow. Despite providing encouraging results in static conditions for objects moving in the detected area, its main limitations reside in the inability to detect objects at ground level and in the reduced 64° field of view.

The NAVITON system [192, 193] processes stereo images to segment out key elements for auditory presentation. For each segmented element, the sonification approach uses discrete pitched sounds, whose pitch, loudness, and temporal delay (depth scanning) depend on object distance, and whose duration is proportional to the depth of the object. Sounds are spatialized with individual HRTFs, custom measured in the full azimuth range and in the vertical plane from −54° to 90°, in 5° steps. Sounds were provided through high-quality open-air reference headphones without headphone compensation. Ten blindfolded participants reported their auditory perception about the sonified virtual 3D scenes in a virtual reality trial, proving to be capable of grasping the general spatial structure of the environment and accurately estimate scene layouts. A real-world navigation scenario was also tested with 5 blind and 5 blindfolded volunteers, who could accurately estimate the spatial position of single obstacles or pairs of obstacles and walk through simple obstacle courses.

The *NAVIG (Navigation Assisted by Artificial VIsion and GNSS)* system [194, 195] aims to enhance mobility and orientation, navigation, object localization, and grasping, both indoors and outdoors. It uses a Global Navigation Satellite System (GNSS) and a rapid visual recognition algorithm. Navigation is ensured by real-time nonindividual HRTF-based rendering, text-to-speech, and semantic sonification metaphors that provide information about the trajectory, position, and the important landmarks in the environment. The 3D audio scenes are conveyed through a bone-conduction headset whose complex frequency response is equalized in order to properly render all the spectral cues of the HRTF. Preliminary experiments have shown that it is possible to design a wearable device that can provide fully analyzed information to the user. However, thorough evaluations of the NAVIG prototype have not been published yet.

5.2. Discussion and Conclusions. The use of HRTFs to code directional information in the above summarized ETAs suggests the importance of a high-fidelity spatial auditory representation of the environment for blind users. However, most of the above works fail to address the hardware- and/or software-related aspects we discussed in Sections 3 and 4, presenting results of performance and usability tests that are based on binaural audio rendering setups that either are ideal yet unrealistic (e.g., [186]) or underestimate the potential of spatial sound itself (e.g., [190]).

As a matter of fact, the preferred choice for the virtual auditory display within the 8 listed ETAs is either individually measured HRTFs or nonindividual, generic HRTFs. Only the cross-modal ETA [187] proposes the use of structural HRTF modeling as a trade-off between localization accuracy and measurement cost. As a result, the evaluation of these systems (often performed through proper localization performance tests) is based either on the best scoring yet unfeasible solution (individually measured HRTFs) or on a costless yet inaccurate one (generic HRTFs), overlooking important aspects in the fidelity of the virtual auditory display such as elevation accuracy and front/back confusion avoidance. Furthermore, the aforementioned monaural localization ability by visually impaired persons (especially early blind) suggests the use of individual pinna cues for azimuth perception, which would make a visually impaired person more vulnerable to degraded localization from nonindividual HRTFs than a sighted person.

Even more unfortunately, the headphones chosen for these tests were in the majority of cases classic circumaural or in-ear headphones that block environmental sounds and thus, as discussed before, are not acceptable for the visually impaired community. The use of a bone-conduction headset is reported only for the SWAN and NAVIG systems [188, 194], where the importance of headphone equalization, although forced to be nonindividual, is also stressed. None of the remaining works, except one [186], even mentions headphone equalization. Effective externalization of the virtual sounds provided to the users is therefore questionable.

It is difficult to rank the importance of the various factors influencing a satisfactory virtual acoustic experience (e.g., externalization, localization accuracy, and front-back confusion rate). Most studies check for only one or two factors and can confirm their influence on one or more spatial sound

perception parameters. Besides the choice of the HRTF set, headphone type, and equalization, and type of sound source (frequency content, familiar/unfamiliar sound, and temporal aspects) [16, 44, 196], other important factors have to be considered. For instance, as explained in Section 2.4, rendering environmental reflections increases externalization, as well as the use of a proper head-tracking method, which also helps in resolving front/back confusion [95]. This may be why most of the above cited studies chose to use high-quality headphones with generic or individual HRTFs, without applying headphone equalization as long as head-tracking or real-time obstacle tracking is implemented. It is also relevant to notice that those systems that use head-mounted cameras to render sounds at locations relative to current head orientation do not even strictly require head-tracking to work dynamically [197].

We believe there is ample space for applying the technologies presented in this review paper to the case of ETAs for the blind. Basic research in HRTF customization techniques is currently in a prolific stage, thanks to advances in computational power and the widespread availability of technologies such as 3D scanning and printing allowing researchers to investigate in detail the relation between individual anthropometry and HRTFs. Although a full and thorough understanding of the mechanisms involved in spatial sound perception still has to be reached, techniques such as HRTF selection, structural HRTF modeling, or HRTF simulations are expected to progressively bridge the gap between accessibility and accuracy of individual binaural audio.

Still it has to be noted that many experiments proved that subjective training to nonindividual HRTFs, especially through cross-modal and game-based training methods, can significantly reduce localization errors in both free field and virtual listening conditions [198]. Feedback can be provided through visual stimuli [199, 200], proprioceptive cues [201, 202], or haptic information [203]. Reductions in front-back confusion rates as large as 40% were reported, as well as improvements in sound localization accuracy in the horizontal and vertical planes regardless of head movement.

On the other hand, the headphone technologies discussed in Section 4 are expected to reach widespread popularity in the blind community. Bone-conduction and active headsets are growing in the consumer market thanks to their affordable price. External multispeaker headsets are still at a prototype stage but from a research point of view open the attractive possibility of introducing individualized binaural playback without the need of fully individual HRTFs. Efforts in the design of such headphones have been produced within the Sound of Vision project [160].

A final comment regards the cosmetic acceptability of the playback device. While bone-conduction and binaural headsets are relatively discreet and portable, external multispeaker headsets may require a bulky and unconventional design. There is considerable variation within the blind community when assessing the cosmetic acceptability of a wearable electronic device, even if it works well. Nevertheless, the visually impaired participants to the survey by Golledge et al. [151] showed overwhelming support for the idea of traveling more often with such a device, independently of its appearance.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement no. 643636.

References

[1] B. D. Simpson, D. S. Brungart, R. C. Dallman, J. Joffrion, M. D. Presnar, and R. H. Gilkey, "Spatial audio as a navigation aid and attitude indicator," in *Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society, HFES '05*, vol. 49, pp. 1602–1606, September 2005.

[2] F. Avanzini, S. Spagnol, A. Rodá, and A. De Götzen, "Designing interactive sound for motor rehabilitation tasks," in *Sonic Interaction Design*, K. Franinovic and S. Serafin, Eds., chapter 12, pp. 273–283, MIT Press, Massachusetts, Mass, USA, 2013.

[3] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 1, pp. 25–35, 2010.

[4] A. Bhowmick and S. M. Hazarika, "An insight into assistive technology for the visually impaired and blind people: State-of-the-art and future trends," *Journal on Multimodal User Interfaces*, vol. 2017, no. 2, pp. 1–24, 2017.

[5] D. I. Ahlmark, *Haptic navigation aids for the visually impaired [Ph.D. thesis]*, Lulea University of Technology, Lulea, Sweden, 2016.

[6] M. Bujacz and P. Strumiłło, "Sonification: review of auditory display solutions in electronic travel aids for the blind," *Archives of Acoustics*, vol. 41, no. 3, pp. 401–414, 2016.

[7] Á. Csapó, G. Wersényi, H. Nagy, and T. Stockman, "A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research," *Journal on Multimodal User Interfaces*, vol. 9, no. 4, pp. 275–286, 2015.

[8] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: effects of beacon sound, capture radius, and practice," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 48, no. 2, pp. 265–278, 2006.

[9] D. S. Brungart, "Near-field virtual audio displays," *Presence: Teleoperators and Virtual Environments*, vol. 11, no. 1, pp. 93–106, 2002.

[10] Á. Csapó and G. Wersényi, "Overview of auditory representations in human-machine interfaces," *ACM Computing Surveys*, vol. 46, no. 2, pp. 1–23, 2013.

[11] Á. Kristjánsson, A. Moldoveanu, Ó. I. Jóhannesson et al., "Designing sensory-substitution devices: principles, pitfalls and potential," *Restorative Neurology and Neuroscience*, vol. 34, no. 5, pp. 769–787, 2016.

[12] J. M. Loomis, J. R. Marston, R. G. Golledge, and R. L. Klatzky, "Personal guidance system for people with visual impairment: a comparison of spatial displays for route guidance," *Journal of Visual Impairment & Blindness*, vol. 99, no. 4, pp. 219–232, 2005.

[13] P. Skulimowski, P. Korbel, and P. Wawrzyniak, "POI explorer - A sonified mobile application aiding the visually impaired in

urban navigation," in *Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, FedCSIS '14*, pp. 969–976, September 2014.

[14] A. Garcia, V. Finomore, G. Burnett, A. Calvo, C. Baldwin, and C. Brill, "Evaluation of multimodal displays for waypoint navigation," in *Proceedings of the 2012 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support, CogSIMA '12*, pp. 134–137, IEEE, Louisiana, La, USA, March 2012.

[15] P. Strumiłło, M. Bujacz, P. Baranski et al., "Different approaches to aiding blind persons in mobility and navigation in the Naviton and Sound of Vision projects," in *Mobility of Visually Impaired People*, E. Pissaloux and R. Velazquez, Eds., pp. 435–468, Springer International Publishing, Cham, Switzerland, 2018.

[16] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, MIT Press, Massachusetts, Mass, USA, 2nd edition, 1996.

[17] S. Spagnol, *Techniques for customized binaural audio rendering with applications to virtual rehabilitation [Ph.D. thesis]*, University of Padova, Padova, Italy, 2012.

[18] J. W. Strutt, "On our perception of sound direction," *Philosophical Magazine*, vol. 13, no. 1907, pp. 214–232, 1907.

[19] G. F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *The Journal of the Acoustical Society of America*, vol. 62, no. 1, pp. 157–167, 1977.

[20] D. S. Brungart, N. I. Durlach, and W. M. Rabinowitz, "Auditory localization of nearby sources. II. Localization of a broadband source," *The Journal of the Acoustical Society of America*, vol. 106, no. 4, pp. 1956–1968, 1999.

[21] A. Wilska, *Studies on Directional Hearing, English translation, Aalto University School of Science and Technology, Department of Signal Processing and Acoustics [Ph.D. thesis]*, University of Helsinki, Helsinki, Finland, 2010.

[22] J. Hebrank and D. Wright, "Spectral cues used in the localization of sound sources on the median plane," *The Journal of the Acoustical Society of America*, vol. 56, no. 6, pp. 1829–1834, 1974.

[23] F. Asano, Y. Suzuki, and T. Sone, "Role of spectral cues in median plane localization," *The Journal of the Acoustical Society of America*, vol. 88, no. 1, pp. 159–168, 1990.

[24] M. B. Gardner and R. S. Gardner, "Problem of localization in the median plane: effect of pinnae cavity occlusion," *The Journal of the Acoustical Society of America*, vol. 53, no. 2, pp. 400–408, 1973.

[25] M. Morimoto, "The contribution of two ears to the perception of vertical angle in sagittal planes," *The Journal of the Acoustical Society of America*, vol. 109, no. 4, pp. 1596–1603, 2001.

[26] J. Hebrank and D. Wright, "Are two ears necessary for localization of sound sources on the median plane?" *The Journal of the Acoustical Society of America*, vol. 56, no. 3, pp. 935–938, 1974.

[27] E. A. G. Shaw and R. Teranishi, "Sound pressure generated in an external-ear replica and real human ears by a nearby point source," *The Journal of the Acoustical Society of America*, vol. 44, no. 1, pp. 240–249, 1968.

[28] S. Spagnol, M. Hiipakka, and V. Pulkki, "A single-azimuth Pinna-Related Transfer Function database," in *Proceedings of the 14th International Conference on Digital Audio Effects, DAFx '11*, pp. 209–212, September 2011.

[29] E. A. Lopez-Poveda and R. Meddis, "A physical model of sound diffraction and reflections in the human concha," *The Journal of the Acoustical Society of America*, vol. 100, no. 5, pp. 3248–3259, 1996.

[30] D. S. Brungart and W. M. Rabinowitz, "Auditory localization of nearby sources. Head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1465–1479, 1999.

[31] S. Spagnol, "On distance dependence of pinna spectral patterns in head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 137, no. 1, pp. EL58–EL64, 2015.

[32] E. A. G. Shaw, "Acoustical features of human ear," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Eds., pp. 25–47, Lawrence Erlbaum Associates, New Jersey, NJ, USA, 1997.

[33] B. C. J. Moore, S. R. Oldfield, and G. J. Dooley, "Detection and discrimination of spectral peaks and notches at 1 and 8 khz," *The Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 820–836, 1989.

[34] K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, "Median plane localization using a parametric model of the head-related transfer function based on spectral cues," *Applied Acoustics*, vol. 68, no. 8, pp. 835–850, 2007.

[35] R. Greff and B. F. G. Katz, "Perceptual evaluation of HRTF notches versus peaks for vertical localisation," in *Proceedings of the 19th International Congress on Acoustics*, 2007.

[36] K. Iida and Y. Ishii, "Roles of spectral peaks and notches in the head-related transfer functions in the upper median plane for vertical localization," *The Journal of the Acoustical Society of America*, vol. 140, no. 4, pp. 2957-2957, 2016.

[37] C. P. Brown and R. O. Duda, "A structural model for binaural sound synthesis," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 6, no. 5, pp. 476–488, 1998.

[38] V. R. Algazi, C. Avendano, and R. O. Duda, "Elevation localization and head-related transfer function analysis at low frequencies," *The Journal of the Acoustical Society of America*, vol. 109, no. 3, pp. 1110–1122, 2001.

[39] O. Kirkeby, E. T. Seppala, A. Karkkainen, L. Karkkainen, and T. Huttunen, "Some effects of the torso on head-related transfer functions," in *Proceedings of the 122nd Audio Engineering Society Convention*, pp. 1045–1052, May 2007.

[40] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, "Auditory distance perception in humans: a summary of past and present research," *Acta Acustica United with Acustica*, vol. 91, no. 3, pp. 409–420, 2005.

[41] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia*, Academic Press Professional, Inc., Massachusetts, Mass, USA, 1994.

[42] M. B. Gardner, "Distance estimation of 0∘ or apparent 0∘-oriented speech signals in anechoic space," *The Journal of the Acoustical Society of America*, vol. 45, no. 1, pp. 47–53, 1969.

[43] D. H. Mershon and J. N. Bowers, "Absolute and relative cues for the auditory perception of egocentric distance," *Perception*, vol. 8, no. 3, pp. 311–322, 1979.

[44] D. S. Brungart, "Auditory localization of nearby sources. III. Stimulus effects," *The Journal of the Acoustical Society of America*, vol. 106, no. 6, pp. 3589–3602, 1999.

[45] F. L. Wightman and D. J. Kistler, "Resolution of front-back ambiguity in spatial hearing by listener and source movement," *The Journal of the Acoustical Society of America*, vol. 105, no. 5, pp. 2841–2853, 1999.

[46] W. R. Thurlow and P. S. Runge, "Effect of induced head movements on localization of direction of sounds," *The Journal of the Acoustical Society of America*, vol. 42, no. 2, pp. 480–488, 1967.

[47] J. M. Speigle and J. M. Loomis, "Auditory distance perception by translating observers," in *Proceedings of the IEEE Research Properties in Virtual Reality Symposium*, pp. 92–99, IEEE, California, Calif, USA, 1993.

[48] M. Rebillat, X. Boutillon, E. T. Corteel, and B. F. G. Katz, "Audio, visual, and audio-visual egocentric distance perception by moving subjects in virtual environments," *ACM Transactions on Applied Perception*, vol. 9, no. 4, article no. 19, 2012.

[49] W. M. Hartmann and A. Wittenberg, "On the externalization of sound images," *The Journal of the Acoustical Society of America*, vol. 99, no. 6, pp. 3678–3688, 1996.

[50] G. Plenge, "On the differences between localization and lateralization," *The Journal of the Acoustical Society of America*, vol. 56, no. 3, pp. 944–951, 1974.

[51] F. L. Wightman and D. J. Kistler, "Headphone simulation of free-field listening. II: psychophysical validation," *The Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 868–878, 1989.

[52] L. S. R. Simon, N. Zacharov, and B. F. G. Katz, "Perceptual attributes for the comparison of head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 140, no. 5, pp. 3623–3632, 2016.

[53] D. R. Begault and E. M. Wenzel, "Headphone localization of speech," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 35, no. 2, pp. 361–376, 1993.

[54] G. Wersényi, "Effect of emulated head-tracking for reducing localization errors in virtual audio simulation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 2, pp. 247–252, 2009.

[55] E. Hendrickx, P. Stitt, J. Messonnier, J. Lyzwa, B. F. Katz, and C. de Boishéraud, "Influence of head tracking on the externalization of speech stimuli for non-individualized binaural synthesis," *The Journal of the Acoustical Society of America*, vol. 141, no. 3, pp. 2011–2023, 2017.

[56] V. Välimäki, J. D. Parker, L. Savioja, J. O. Smith, and J. S. Abel, "Fifty years of artificial reverberation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 20, no. 5, pp. 1421–1448, 2012.

[57] Y. Yuan, L. Xie, Z.-H. Fu, M. Xu, and Q. Cong, "Sound image externalization for headphone based real-time 3D audio," *Frontiers of Computer Science*, vol. 11, no. 3, pp. 419–428, 2017.

[58] S. Werner, G. Götz, and F. Klein, "Influence of head tracking on the externalization of auditory events at divergence between synthesized and listening room using a binaural headphone system," in *Proceedings of the 142nd Audio Engineering Society International Convention*, 2017.

[59] L. H. Benedetti and M. Loeb, "A comparison of auditory monitoring performance in blind subjects with that of sighted subjects in light and dark," *Perception & Psychophysics*, vol. 11, no. 1, pp. 10–16, 1972.

[60] I. Starlinger and W. Niemeyer, "Do the blind hear better? investigations on auditory processing in congenital or early acquired blindness I. Peripheral functions," *International Journal of Audiology*, vol. 20, no. 6, pp. 503–509, 1981.

[61] M. Bross and M. Borenstein, "Temporal auditory acuity in blind and sighted subjects: a signal detection analysis," *Perceptual and Motor Skills*, vol. 55, no. 3, pp. 963–966, 1982.

[62] H.-H. Lai and Y.-C. Chen, "A study on the blind's sensory ability," *International Journal of Industrial Ergonomics*, vol. 36, no. 6, pp. 565–570, 2006.

[63] W. Niemeyer and I. Starlinger, "Do the blind hear better? investigations on auditory processing in congenital or early acquired blindness II. Central functions," *International Journal of Audiology*, vol. 20, no. 6, pp. 510–515, 1981.

[64] C. Muchnik, M. Efrati, E. Nemeth, M. Malin, and M. Hildesheimer, "Central auditory skills in blind and sighted subjects," *Scandinavian Audiology*, vol. 20, no. 1, pp. 19–23, 1991.

[65] J. P. Rauschecker, "Compensatory plasticity and sensory substitution in the cerebral cortex," *Trends in Neurosciences*, vol. 18, no. 1, pp. 36–43, 1995.

[66] N. Lessard, M. Paré, F. Lepore, and M. Lassonde, "Early-blind human subjects localize sound sources better than sighted subjects," *Nature*, vol. 395, no. 6699, pp. 278–280, 1998.

[67] M. P. Zwiers, A. J. Van Opstal, and J. R. Cruysberg, "A spatial hearing deficit in early-blind humans," *The Journal of Neuroscience*, vol. 21, no. 9, pp. 1–5, 2001.

[68] M. Ohuchi, Y. Iwaya, Y. Suzuki, and T. Munekata, "A comparative study of sound localization acuity of congenital blind and sighted people," *Acoustical Science and Technology*, vol. 27, no. 5, pp. 290–293, 2006.

[69] A. J. Kolarik, S. Pardhan, S. Cirstea, and B. C. J. Moore, "Auditory spatial representations of the world are compressed in blind humans," *Experimental Brain Research*, vol. 235, no. 2, pp. 597–606, 2017.

[70] C. E. Rice, "Human echo perception," *Science*, vol. 155, no. 3763, pp. 656–664, 1967.

[71] P. Voss, V. Tabry, and R. J. Zatorre, "Trade-off in the sound localization abilities of early blind individuals between the horizontal and vertical planes," *The Journal of Neuroscience*, vol. 35, no. 15, pp. 6051–6056, 2015.

[72] B. Röder, W. Teder-Salejarvi, A. Sterr, F. Rosler, S. A. Hillyard, and H. J. Neville, "Improved auditory spatial tuning in blind humans," *Nature*, vol. 400, no. 6740, pp. 162–166, 1999.

[73] B. J. Cratty, *Movement and Spatial Awareness in Blind Children and Youth*, C. C. Thomas, Ed., Springfield, Illinois, Ill, USA, 1971.

[74] A. Dufour, O. Després, and V. Candas, "Enhanced sensitivity to echo cues in blind subjects," *Experimental Brain Research*, vol. 165, no. 4, pp. 515–519, 2005.

[75] I. G. Bassett and E. J. Eastmond, "Echolocation: measurement of pitch versus distance for sounds reflected from a flat surface," *The Journal of the Acoustical Society of America*, vol. 36, no. 5, pp. 911–916, 1964.

[76] L. D. Rosenblum, M. S. Gordon, and L. Jarquin, "Echolocating distance by moving and stationary listeners," *Ecological Psychology Journal*, vol. 12, no. 3, pp. 181–206, 2000.

[77] W. N. Kellogg, "Sonar System of the Blind," *Science*, vol. 137, no. 3528, pp. 399–404, 1962.

[78] T. Miura, T. Muraoka, and T. Ifukube, "Comparison of obstacle sense ability between the blind and the sighted: A basic psychophysical study for designs of acoustic assistive devices," *Acoustical Science and Technology*, vol. 31, no. 2, pp. 137–147, 2010.

[79] C. E. Rice and S. H. Feinstein, "Sonar system of the blind: size discrimination," *Science*, vol. 38, no. 148, pp. 1107-1108, 1965.

[80] T. A. Stoffregen and J. B. Pittenger, "Human echolocation as a basic form of perception and action," *Ecological Psychology Journal*, vol. 7, no. 3, pp. 181–216, 1995.

[81] G. Wersényi, "Virtual localization by blind persons," *Journal of the Audio Engineering Society*, vol. 60, no. 7-8, pp. 568–579, 2012.

[82] A. Dobrucki, P. Plaskota, P. Pruchnicki, M. Pec, M. Bujacz, and P. Strumiłło, "Measurement system for personalized head-related transfer functions and its verification by virtual source

localization trials with visually impaired and sighted individuals," *Journal of the Audio Engineering Society*, vol. 58, no. 9, pp. 724–738, 2010.

[83] A. Afonso, A. Blum, B. F. G. Katz, P. E. Tarroux, G. Borst, and M. Denis, "Structural properties of spatial representations in blind people: scanning images constructed from haptic exploration or from locomotion in a 3-D audio virtual environment," *Memory & Cognition*, vol. 38, no. 5, pp. 591–604, 2010.

[84] L. Picinali, A. Afonso, M. Denis, and B. F. G. Katz, "Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge," *International Journal of Human-Computer Studies*, vol. 72, no. 4, pp. 393–407, 2014.

[85] A. Cobo, N. E. Guerrón, C. Martín, F. del Pozo, and J. J. Serrano, "Differences between blind people's cognitive maps after proximity and distant exploration of virtual environments," *Computers in Human Behavior*, vol. 77, no. 12, pp. 294–308, 2017.

[86] K. Papadopoulos, E. Koustriava, and M. Barouti, "Cognitive maps of individuals with blindness for familiar and unfamiliar spaces: construction through audio-tactile maps and walked experience," *Computers in Human Behavior*, vol. 75, no. 10, pp. 376–384, 2017.

[87] C. I. Cheng and G. H. Wakefield, "Introduction to head-related transfer functions (HRTFs): Representations of HRTFs in time, frequency, and space," *Journal of the Audio Engineering Society*, vol. 49, no. 4, pp. 231–249, 2001.

[88] J. Blauert, *The Technology of Binaural Listening*, Springer, New York, NY, USA, 2013.

[89] A. W. Bronkhorst, "Localization of real and virtual sound sources," *The Journal of the Acoustical Society of America*, vol. 98, no. 5, pp. 2542–2553, 1995.

[90] G. Wersényi, "Localization in a head-related transfer function-based virtual audio synthesis using additional high-pass and low-pass filtering of sound sources," *Acoustical Science and Technology*, vol. 28, no. 4, pp. 244–250, 2007.

[91] M. D. Burkhard and R. M. Sachs, "Anthropometric manikin for acoustic research," *The Journal of the Acoustical Society of America*, vol. 58, no. 1, pp. 214–222, 1975.

[92] A. Andreopoulou, D. R. Begault, and B. F. G. Katz, "Inter-laboratory round robin HRTF measurement comparison," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 5, pp. 895–906, 2015.

[93] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Proceedings of the IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pp. 99–102, IEEE, New York, NY, USA, 2001.

[94] E. M. Wenzel, M. Arruda, D. J. Kistler, and F. L. Wightman, "Localization using nonindividualized head-related transfer functions," *The Journal of the Acoustical Society of America*, vol. 94, no. 1, pp. 111–123, 1993.

[95] D. R. Begault, E. M. Wenzel, and M. R. Anderson, "Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source," *Journal of the Audio Engineering Society*, vol. 49, no. 10, pp. 904–916, 2001.

[96] H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi, "Binaural technique: do we need individual recordings?" *Journal of the Audio Engineering Society*, vol. 44, no. 6, pp. 451–464, 1996.

[97] M. Geronazzo, S. Spagnol, and F. Avanzini, "A modular framework for the analysis and synthesis of head-related transfer functions," in *Proceedings of the 134th AES Convention - Audio Engineering Society*, 2013.

[98] B. U. Seeber and H. Fastl, "Subjective selection of non-individual head-related transfer functions," in *Proceedings of the International Conference on Auditory Display (ICAD '03)*, pp. 259–262, 2003.

[99] D. N. Zotkin, R. Duraiswami, and L. S. Davis, "Rendering localized spatial audio in a virtual auditory space," *IEEE Transactions on Multimedia*, vol. 6, no. 4, pp. 553–564, 2004.

[100] M. Geronazzo, S. Spagnol, A. Bedin, and F. Avanzini, "Enhancing vertical localization with image-guided selection of non-individual head-related transfer functions," in *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '14*, pp. 4496–4500, May 2014.

[101] K. Iida, Y. Ishii, and S. Nishioka, "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," *The Journal of the Acoustical Society of America*, vol. 136, no. 1, pp. 317–333, 2014.

[102] R. H. Y. So, B. Ngan, A. Horner, J. Braasch, J. Blauert, and K. L. Leung, "Toward orthogonal non-individualised head-related transfer functions for forward and backward directional sound: Cluster analysis and an experimental study," *Ergonomics*, vol. 53, no. 6, pp. 767–781, 2010.

[103] B. F. G. Katz and G. Parseihian, "Perceptually based head-related transfer function database optimization," *The Journal of the Acoustical Society of America*, vol. 131, no. 2, pp. EL99–EL105, 2012.

[104] S. Hwang, Y. Park, and Y.-S. Park, "Modeling and customization of head-related impulse responses based on general basis functions in time domain," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 965–980, 2008.

[105] K. H. Shin and Y. Park, "Enhanced vertical perception through head-related impulse response customization based on pinna response tuning in the median plane," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E91-A, no. 1, pp. 345–356, 2008.

[106] W. M. Rabinowitz, J. Maxwell, Y. Shao, and M. Wei, "Sound localization cues for a magnified head: implications from sound diffraction about a rigid sphere," *Presence: Teleoperators and Virtual Environments*, vol. 2, no. 2, pp. 125–129, 1993.

[107] S. Spagnol, E. Tavazzi, and F. Avanzini, "Distance rendering and perception of nearby virtual sound sources with a near-field filter model," *Applied Acoustics*, vol. 115, pp. 61–73, 2017.

[108] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Acoustic simulation of KEMAR's HRTFs: verification with measurements and the effects of modifying head shape and pinna concavity," in *Proceeding of the Interenational Workshop Principles Apply Spatial Hearing (IWPASH '09)*, 2009.

[109] R. O. Duda, C. Avendano, and V. R. Algazi, "An adaptable ellipsoidal head model for the interaural time difference," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 965–968, March 1999.

[110] R. Bomhardt, M. Lins, and J. Fels, "Analytical ellipsoidal model of interaural time differences for the individualization of head-related impulse responses," *Journal of the Audio Engineering Society*, vol. 64, no. 11, pp. 882–894, 2016.

[111] B. F. G. Katz and M. Noisternig, "A comparative study of interaural time delay estimation methods," *The Journal of the Acoustical Society of America*, vol. 135, no. 6, pp. 3530–3540, 2014.

[112] A. Andreopoulou and B. F. Katz, "Identification of perceptually relevant methods of inter-aural time difference estimation," *The*

Journal of the Acoustical Society of America, vol. 142, no. 2, pp. 588–598, 2017.

[113] V. R. Algazi, R. O. Duda, and D. M. Thompson, "The use of head-and-torso models for improved spatial sound synthesis in," in *Proceedings of the 113th Convention of the Audio Engineering Society*, pp. 1–18, 2002.

[114] V. R. Algazi, R. O. Duda, R. Duraiswami, N. A. Gumerov, and Z. Tang, "Approximating the head-related transfer function using simple geometric models of the head and torso," *The Journal of the Acoustical Society of America*, vol. 112, no. 5, pp. 2053–2064, 2002.

[115] A. Meshram, R. Mehra, H. Yang, E. Dunn, J.-M. Franm, and D. Manocha, "P-HRTF: Efficient personalized HRTF computation for high-fidelity spatial sound," in *Proceedings of the 13th IEEE International Symposium on Mixed and Augmented Reality, ISMAR '14*, pp. 53–61, IEEE, Munich, Germany, September 2014.

[116] V. Algazi, R. Duda, R. Morrison, and D. Thompson, "Structural composition and decomposition of HRTFs," in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 103–106, IEEE, New York, NY, USA, 2001.

[117] V. R. Algazi, C. Avendano, and R. O. Duda, "Estimation of a spherical-head model from anthropometry," *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 472–479, 2001.

[118] S. Spagnol and F. Avanzini, "Anthropometric tuning of a spherical head model for binaural virtual acoustics based on interaural level differences," in *Proceedings of the 21st International Conference on Auditory Display ICAD '15*, pp. 204–209, 2015.

[119] M. Aussal, F. Alouges, and B. F. Katz, "HRTF interpolation and ITD personalization for binaural synthesis using spherical harmonics," in *Proceedings of the 25th UK Conference Audio Engineering Society*, 2012.

[120] A. J. Watkins, "Psychoacoustical aspects of synthesized vertical locale cues," *The Journal of the Acoustical Society of America*, vol. 63, no. 4, pp. 1152–1165, 1978.

[121] R. Teranishi and E. A. G. Shaw, "External-ear acoustic models with simple geometry," *The Journal of the Acoustical Society of America*, vol. 44, no. 1, pp. 257–263, 1968.

[122] E. A. G. Shaw, "The acoustics of the external ear," in *Acoustical Factors Affecting Hearing Aid Performance*, G. A. Studebaker and I. Hochberg, Eds., pp. 109–125, University Park Press, Maryland, Md, USA, 1980.

[123] P. Satarzadeh, R. V. Algazi, and R. O. Duda, "Physical and filter pinna models based on anthropometry," in *Proceedings of the 122nd Audio Engineering Society Convention '07*, pp. 718–737, May 2007.

[124] K. J. Faller, A. Barreto, and M. Adjouadi, "Augmented Hankel total least-squares decomposition of head-related transfer functions," *Journal of the Audio Engineering Society*, vol. 58, no. 1-2, pp. 3–21, 2010.

[125] S. Spagnol, M. Geronazzo, and F. Avanzini, "On the relation between pinna reflection patterns and head-related transfer function features," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 21, no. 3, pp. 508–520, 2013.

[126] L. Li and Q. Huang, "HRTF personalization modeling based on RBF neural network," in *Proceedings of the 2013 38th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '13*, pp. 3707–3710, IEEE, Vancouver, BC, Canada, May 2013.

[127] Q. Huang and L. Li, "Modeling individual HRTF tensor using high-order partial least squares," *Journal on Advances in Signal Processing*, vol. 58, pp. 1–14, 2014.

[128] F. Grijalva, L. Martini, S. Goldenstein, and D. Florencio, "Anthropometric-based customization of head-related transfer functions using Isomap in the horizontal plane," in *Proceedings of the 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '14*, pp. 4506–4510, IEEE, Firenze, Italy, May 2014.

[129] P. Bilinski, J. Ahrens, M. R. P. Thomas, I. J. Tashev, and J. C. Platt, "HRTF magnitude synthesis via sparse representation of anthropometric features," in *Proceedings of the 39th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '14*, pp. 4501–4505, IEEE, Firenze, Italy, May 2014.

[130] J. C. Middlebrooks, "Individual differences in external-ear transfer functions reduced by scaling in frequency," *The Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1480–1492, 1999.

[131] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Frequency and amplitude estimation of the first peak of head-related transfer functions from individual pinna anthropometry," *The Journal of the Acoustical Society of America*, vol. 137, no. 2, pp. 690–701, 2015.

[132] P. Mokhtari, H. Takemoto, R. Nishimura, and H. Kato, "Vertical normal modes of human ears: Individual variation and frequency estimation from pinna anthropometry," *The Journal of the Acoustical Society of America*, vol. 140, no. 2, pp. 814–831, 2016.

[133] V. C. Raykar, R. Duraiswami, and B. Yegnanarayana, "Extracting the frequencies of the pinna spectral notches in measured head related impulse responses," *The Journal of the Acoustical Society of America*, vol. 118, no. 1, pp. 364–374, 2005.

[134] S. Spagnol and F. Avanzini, "Frequency estimation of the first pinna notch in head-related transfer functions with a linear anthropometric model," in *Proceedings of the 18th International Conference on Digital Audio Effects*, pp. 231–236, 2015.

[135] S. Spagnol, M. Geronazzo, D. Rocchesso, and F. Avanzini, "Synthetic individual binaural audio delivery by pinna image processing," *International Journal of Pervasive Computing and Communications*, vol. 10, no. 3, pp. 239–254, 2014.

[136] S. Spagnol, S. Scaiella, M. Geronazzo, and F. Avanzini, "Subjective evaluation of a low-order parametric filter model of the pinna for binaural sound rendering," in *Proceedings of the 22nd International Congress on Sound and Vibration, ICSV '15*, July 2015.

[137] T. Huttunen, E. T. Seppälä, O. Kirkeby, A. Kärkkäinen, and L. Kärkkäinen, "Simulation of the transfer function for a head-and-torso model over the entire audible frequency range," *Journal of Computational Acoustics*, vol. 15, no. 4, pp. 429–448, 2007.

[138] B. F. G. Katz, "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation," *The Journal of the Acoustical Society of America*, vol. 110, no. 5, pp. 2440–2448, 2001.

[139] Y. Kahana and P. A. Nelson, "Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric models," *Journal of Sound and Vibration*, vol. 300, no. 3-5, pp. 552–579, 2007.

[140] W. Kreuzer, P. Majdak, and Z. Chen, "Fast multipole boundary element method to calculate head-related transfer functions for a wide frequency range," *The Journal of the Acoustical Society of America*, vol. 126, no. 3, pp. 1280–1290, 2009.

[141] N. A. Gumerov, A. E. O'Donovan, R. Duraiswami, and D. N. Zotkin, "Computation of the head-related transfer function via the fast multipole accelerated boundary element method

and its spherical harmonic representation," *The Journal of the Acoustical Society of America*, vol. 127, no. 1, pp. 370–386, 2010.

[142] H. Ziegelwanger, P. Majdak, and W. Kreuzer, "Numerical calculation of listener-specific head-related transfer functions and sound localization: microphone model and mesh discretization," *The Journal of the Acoustical Society of America*, vol. 138, no. 1, pp. 208–222, 2015.

[143] C. T. Jin, P. Guillon, N. Epain et al., "Creating the Sydney York morphological and acoustic recordings of ears database," *IEEE Transactions on Multimedia*, vol. 16, no. 1, pp. 37–46, 2014.

[144] H. Ziegelwanger, W. Kreuzer, and P. Majdak, "MESH2HRTF: an open-source software package for the numerical calculation of head-related transfer functions," in *Proceedings of the 22nd International Congress on Sound and Vibration 2015 (ICSV '22)*, IEEE, Firenze, Italy, 2015.

[145] M. Dellepiane, N. Pietroni, N. Tsingos, M. Asselot, and R. Scopigno, "Reconstructing head models from photographs for individualized 3D-audio processing," *Computer Graphics Forum*, vol. 27, no. 7, pp. 1719–1727, 2008.

[146] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sørensen, "Transfer characteristics of headphones measured on human ears," *Journal of the Audio Engineering Society*, vol. 43, no. 4, pp. 203–217, 1995.

[147] D. Pralong and S. Carlile, "The role of individualized headphone calibration for the generation of high fidelity virtual auditory space," *The Journal of the Acoustical Society of America*, vol. 100, no. 6, pp. 3785–3793, 1996.

[148] B. Masiero and J. Fels, "Perceptually robust headphone equalization for binaural reproduction," in *Proceedings of the 130th Audio Engineering Society Convention*, 2011.

[149] Z. Schärer and A. Lindau, "Evaluation of equalization methods for binaural signals," in *Proceedings of the 126th Audio Engineering Society Convention*, 2009.

[150] D. Schonstein, L. Ferré, and B. F. Katz, "Comparison of headphones and equalization for virtual auditory source localization," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 3724-3724, 2008.

[151] R. G. Golledge, J. R. Marston, J. M. Loomis, and R. L. Klatzky, "Stated Preferences for Components of a Personal Guidance System for Nonvisual Navigation," *Journal of Visual Impairment & Blindness*, vol. 98, no. 3, pp. 135–147, 2004.

[152] V. Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[153] M. A. Gerzon, "Ambisonics in multichannel broadcasting and video," *Journal of the Audio Engineering Society*, vol. 33, no. 11, pp. 859–871, 1985.

[154] F. M. König, "4-canal headphone for in-front localization and HDTV- or Dolby-surround use," in *Proceedings of the 96th Convention Audio Engineering Society*, 1994.

[155] F. M. König, "A new supra-aural dynamic headphone system for in-front localization and surround reproduction of sound," in *Proceedings of the 102nd Convention Audio Engineering Society*, 1997.

[156] F. M. König, "New measurements and psychoacoustic investigations on a headphone for TAX/HDTV/Dolby-surround reproduction of sound," in *Proceedings of the 98th Convention Audio Engineering Society*, 1995.

[157] S. G. Weinrich, "Improved externalization and frontal perception of headphone signals," in *Proceedings of the 92nd Convention Audio Engineering Society*, 1992.

[158] K. Sunder, E.-L. Tan, and W.-S. Gan, "Individualization of binaural synthesis using frontal projection headphones," *Journal of the Audio Engineering Society*, vol. 61, no. 12, pp. 989–1000, 2013.

[159] R. Greff and B. F. Katz, "Circumaural transducer arrays for binaural synthesis," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 3562-3562, 2008.

[160] M. Bujacz, K. Kropidlowski, G. Ivanica et al., "Sound of Vision - Spatial audio output and sonification approaches," in *Proceedings of the 15th International Conference of Computers Helping People with Special Needs (ICCHP '16)*, K. Miesenberger, C. Bühler, and P. Penaz, Eds., vol. 9759, pp. 202–209, Springer, Linz, Austria, 2016.

[161] B. N. Walker and R. M. Stanley, "Thresholds of audibility for bone-conduction headsets," in *Proceedings of the International Conference on Auditory Display*, pp. 218–222, IEEE, Limerick, Ireland, 2005.

[162] B. N. Walker, R. M. Stanley, N. Iyer, B. D. Simpson, and D. S. Brungart, "Evaluation of bone-conduction headsets for use in multitalker communication environments," in *Proceedings of the 49th Annual Meeting of Human Factors and Ergonomics Society*, pp. 1615–1619, 2005.

[163] R. M. Stanley and B. N. Walker, "Lateralization of sound using bone-conduction headsets," in *Proceedings of the 50th Annual Meeting of Human Factors and Ergonomics Society*, pp. 1571–1575, 2006.

[164] J. A. MacDonald, P. P. Henry, and T. R. Letowski, "Spatial audio through a bone conduction interface," *International Journal of Audiology*, vol. 45, no. 10, pp. 595–599, 2006.

[165] R. M. Stanley, *Measurement and validation of bone-conduction adjustment functions in virtual 3D audio displays [Ph.D. thesis]*, Georgia Institute of Technology, Georgia, Ga, USA, 2009.

[166] S. Reinfeldt, B. Håkansson, H. Taghavi, and M. Eeg-Olofsson, "New developments in bone-conduction hearing implants: a review," *Medical Devices: Evidence and Research*, vol. 8, pp. 79–93, 2015.

[167] B. N. Walker and J. L. Lindsay, "Navigation performance in a virtual environment with bonephones," in *Proceedings of the 11th International Conf.erence Auditory Display (ICAD '05)*, pp. 260–263, IEEE, Limerick, Ireland, 2005.

[168] B. N. Walker, R. M. Stanley, and A. Przekwas, "High fidelity modeling and experimental evaluation of binaural bone conduction communication devices," in *Proceedings of the 19th International Congress on Acoustics*, 2007.

[169] R. W. Lindeman, H. Noma, and P. G. De Barros, "Hear-through and mic-through augmented reality: Using bone conduction to display spatialized audio," in *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR '07*, November 2007.

[170] R. W. Lindeman, H. Noma, and P. G. De Barros, "An empirical study of Hear-Through augmented reality: Using bone conduction to deliver spatialized audio," in *Proceedings of the IEEE Virtual Reality*, pp. 35–42, March 2008.

[171] A. Barde, W. S. Helton, G. Lee, and M. Billinghurst, "Binaural spatialization over a bone conduction headset: Minimum discernable angular difference," in *Proceedings of the 140th Convention of the Audio Engineering Society*, 2016.

[172] B. Rafaely, "Active noise reducing headset - An overview," in *Proceedings of the International Congress and Exhibition on Noise Control Engineering*, 2001.

[173] A. V. Oppenheim, E. Weinstein, K. C. Zangi, M. Feder, and D. Gauger, "Single-sensor active noise cancellation," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 2, no. 2, pp. 285–290, 1994.

[174] M. Tikander, "Usability issues in listening to natural sounds with an augmented reality audio headset," *Journal of the Audio Engineering Society*, vol. 57, no. 6, pp. 430–441, 2009.

[175] A. Härmä, J. Jakka, M. Tikander et al., "Augmented reality audio for mobile and wearable appliances," *Journal of the Audio Engineering Society*, vol. 52, no. 6, pp. 618–639, 2004.

[176] M. Tikander, M. Karjalainen, and V. Riikonen, "An augmented reality audio headset," in *Proceedings of the 11th International Conference on Digital Audio Effects, DAFx '08*, 2008.

[177] J. Liski, *Adaptive hear-through headset [M.S. thesis]*, Aalto University School of Electrical Engineering, Espoo, Finland, 2016.

[178] S. Brunner, H. J. Maempel, and S. Weinzierl, "On the audibility of comb filter distortions," in *Proceedings of the 122nd Convention Audio Engineering Society*, 2007.

[179] J. Rämö and V. Välimäki, "Digital augmented reality audio headset," *Journal of Electrical and Computer Engineering*, vol. 2012, Article ID 457374, 13 pages, 2012.

[180] D. Hammershøi and H. Møller, "Sound transmission to and within the human ear canal," *The Journal of the Acoustical Society of America*, vol. 100, no. 1, pp. 408–427, 1996.

[181] P. B. L. Meijer, "An experimental system for auditory image representations," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 2, pp. 112–121, 1992.

[182] M. A. Hersh and M. A. Johnson, *Assistive Technology for Visually Impaired and Blind People*, Springer, London, UK, 1st edition, 2008.

[183] S. Shoval, J. Borenstein, and Y. Koren, "Auditory guidance with the Navbelt-a computerized travel aid for the blind," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 28, no. 3, pp. 459–467, 1998.

[184] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, and S. Yaacob, "A stereo image processing system for visually impaired," *International Journal of Signal Processing*, vol. 2, no. 3, pp. 136–145, 2008.

[185] F. Zhigang and L. Ting, "Audification-based electronic travel aid system," in *Proceedings of the International Conference on Computer Design and Applications, ICCDA '10*, vol. 5, pp. 137–141, IEEE, Qinhuangdao, China, 2010.

[186] J. L. González-Mora, A. Rodríguez-Hernández, L. F. Rodríguez-Ramos, L. Díaz-Saco, and N. Sosa, "Development of a new space perception system for blind people, based on the creation of a virtual acoustic space," in *Engineering Applications of Bio-Inspired Artificial Neural Networks*, vol. 1607 of *Lecture Notes in Computer Science*, pp. 321–330, Springer, Berlin, Germany, 1999.

[187] F. Fontana, A. Fusiello, M. Gobbi et al., "A cross-modal electronic travel aid device," in *Human Computer Interaction with Mobile Devices*, vol. 2411 of *Lecture Notes in Computer Science*, pp. 393–397, Springer, Berlin, Germany, 2002.

[188] J. Wilson, B. N. Walker, J. Lindsay, C. Cambias, and F. Dellaert, "SWAN: system for wearable audio navigation," in *Proceedings of the 11th IEEE International Symposium on Wearable Computers, ISWC '07*, pp. 91–98, IEEE, Massachusetts, Mass, USA, October 2007.

[189] M. A. Torres-Gil, O. Casanova-Gonzalez, and J. L. Gonzalez-Mora, "Applications of virtual reality for visually impaired people," *WSEAS Transactions on Computers*, vol. 9, no. 2, pp. 184–193, 2010.

[190] L. Dunai, G. P. Fajarnes, V. S. Praderas, B. D. Garcia, and I. L. Lengua, "Real–Time Assistance Prototype –A new navigation aid for blind people," in *Proceedings of the 36th Annual Conference of the IEEE Industrial Electronics Society, IECON '10*, pp. 1173–1178, IEEE, Arizona, Ariz, USA, November 2010.

[191] G. P. Fajarnes, L. Dunai, and V. S. Praderas, "CASBLiP - A new cognitive object detection and orientation system for impaired people," in *Proceedings of the 4th International Conference on Cognitive Systems*, IEEE, Zurich, Switzerland, 2010.

[192] M. Bujacz, P. Skulimowski, and P. Strumiłło, "Sonification of 3D scenes using personalized spatial audio to aid visually impaired persons," in *Proceedings of the 17th International Conference on Auditory Display*, 2011.

[193] M. Bujacz, P. Skulimowski, and P. Strumillo, "Naviton-a prototype mobility aid for auditory presentation of three-dimensional scenes to the visually impaired," *Journal of the Audio Engineering Society*, vol. 60, no. 9, pp. 696–708, 2012.

[194] B. F. G. Katz, S. Kammoun, G. Parseihian et al., "NAVIG: augmented reality guidance system for the visually impaired," *Virtual Reality*, vol. 16, no. 4, pp. 253–269, 2012.

[195] B. F. G. Katz, F. Dramas, G. Parseihian et al., "NAVIG: Guidance system for the visually impaired using virtual augmented reality," *Technology and Disability*, vol. 24, no. 2, pp. 163–178, 2012.

[196] J. Vliegen and A. J. Van Opstal, "The influence of duration and level on human sound localization," *The Journal of the Acoustical Society of America*, vol. 115, no. 4, pp. 1705–1713, 2004.

[197] S. Spagnol, R. Hoffmann, M. Herrera Martínez, and R. Unnthorsson, "Blind wayfinding with physically-based liquid sounds," *International Journal of Human-Computer Studies*, vol. 115, pp. 9–19, 2018.

[198] C. Mendonça, "A review on auditory space adaptations to altered head-related cues," *Frontiers in Neuroscience*, vol. 8, no. 219, Article ID Article 219, pp. 1–14, 2014.

[199] P. Zahorik, P. Bangayan, V. Sundareswaran, K. Wang, and C. Tam, "Perceptual recalibration in human sound localization: Learning to remediate front-back reversals," *The Journal of the Acoustical Society of America*, vol. 120, no. 1, pp. 343–359, 2006.

[200] S. Carlile, "The plastic ear and perceptual relearning in auditory spatial perception," *Frontiers in Neuroscience*, vol. 8, no. 237, pp. 1–13, 2014.

[201] G. Parseihian and B. F. G. Katz, "Rapid head-related transfer function adaptation using a virtual auditory environment.," *The Journal of the Acoustical Society of America*, vol. 131, no. 4, pp. 2948–2957, 2012.

[202] A. Honda, H. Shibata, S. Hidaka, J. Gyoba, Y. Iwaya, and Y. Suzuki, "Effects of head movement and proprioceptive feedback in training of sound localization," *i-Perception*, vol. 4, no. 4, pp. 253–264, 2013.

[203] O. Balan, A. Moldoveanu, H. Nagy et al., "Haptic-auditory perceptual feedback based training for improving the spatial acoustic resolution of the visually impaired people," in *Proceedings of the 21st International Conference on Auditory Display (ICAD '15)*, pp. 21–28, 2015.

Hindawi
Wireless Communications and Mobile Computing
Volume 2018, Article ID 8971206, 9 pages
https://doi.org/10.1155/2018/8971206

Research Article

Cognitive Load Assessment from EEG and Peripheral Biosignals for the Design of Visually Impaired Mobility Aids

Charalampos Saitis,[1] **Mohammad Zavid Parvez**,[2] **and Kyriaki Kalimeri**[2]

[1]*Audio Communication Group, Technical University of Berlin, Berlin, Germany*
[2]*Data Science Lab, Institute for Scientific Interchange (ISI Foundation), Turin, Italy*

Correspondence should be addressed to Charalampos Saitis; charalampos.saitis@campus.tu-berlin.de

Received 23 September 2017; Revised 10 January 2018; Accepted 30 January 2018; Published 28 February 2018

Academic Editor: Tao Gu

Reliable detection of cognitive load would benefit the design of intelligent assistive navigation aids for the visually impaired (VIP). Ten participants with various degrees of sight loss navigated in unfamiliar indoor and outdoor environments, while their electroencephalogram (EEG) and electrodermal activity (EDA) signals were being recorded. In this study, the cognitive load of the tasks was assessed in real time based on a modification of the well-established event-related (de)synchronization (ERD/ERS) index. We present an in-depth analysis of the environments that mostly challenge people from certain categories of sight loss and we present an automatic classification of the perceived difficulty in each time instance, inferred from their biosignals. Given the limited size of our sample, our findings suggest that there are significant differences across the environments for the various categories of sight loss. Moreover, we exploit cross-modal relations predicting the cognitive load in real time inferring on features extracted from the EDA. Such possibility paves the way for the design on less invasive, wearable assistive devices that take into consideration the well-being of the VIP.

1. Introduction

Visual impairment affects approximately 285 million individuals worldwide according to the WHO [1]. Assistive navigation aids are essential to the visually impaired (VIP) for improving their quality of life and increase their independence. Traditionally, VIP relied exclusively on the white cane due to its simplicity; despite its reliability in obstacle detection, it does not provide any information regarding important aspects of navigation such as the distance, the speed, or the shortest path to the destination [2]. New technologies came to fill this gap, enhancing the traditional assistive aids, aiming to improve the route planning [3], navigating long distances [4], discovering landmarks [5], and detecting obstacles [6–8]. Ranging from smartphone applications to wearable devices, assistive navigation aids promote greater independence and enable VIP to perform tasks formerly impossible or difficult to accomplish [9]. Yet, the focus of these aids is often on optimizing way-finding or localization tasks without taking into consideration the individual's needs [10].

Building on our previous work [11, 12], in this study, we place the focus entirely on the visually impaired, assessing biomarkers that can predict in real time the mental effort of the visually impaired while navigating in unfamiliar indoor and outdoor urban environments. The challenges VIP experience during orientation and mobility tasks can be framed according to the cognitive load theory [13], since, during orientation and navigation tasks, a specific amount of space is consumed by the working memory for the exact cognitive demands necessary.

We designed two ad hoc orientation and mobility tasks gathering a wide range of behavioral and biophysical signals from 10 VIP with various categories of sight loss (see Table 1), who volunteered to participate in our study. Collecting electroencephalogram (EEG) signals, we assess the cognitive load and task engagement in the performed task. EEG signals are shown to be stable indicators of the cognitive load in a

Table 1: Category of vision impairment and gender of participants.

Category	Description	Outdoor route	Indoor route
VI-2	Vision is less than 10% and more than 5%	1 (F)	2 (F, M)
VI-3	Vision is less than 5% and more than being able to count fingers less than one meter away	3 (F, F, M)	4 (F, F, M, F)
VI-4	Not being able to count fingers less than one meter away	2 (F, M)	3 (F, M, F)
VI-5	No light perception	1 (F)	–

Based on the classi_cation of visual impairment by the World Health Organization (http://apps.who.int/classifications/icd10/browse/2016/en/#/H53-H54).

Table 2: Descriptions and mobility challenges of the different indoor and outdoor scenes.

Route	ID	Scene	Challenges
Outdoor	A	Shopping street	People, ads, chairs, tables, poles
	B	Small street	People, poles, ads
	C	Narrow alley	People, chairs, tables, street ads, trash bins
	D	Urban park	People
	E	Open space	People
	F	Crossing road	People
	G	Crossing street	People
	H	construction	People
Indoor	A	Door	Automated doors (hinged and rotating)
	B	Elevator	Calling the elevator, selecting floor
	C	Corridor	People with noise, doors open suddenly
	D	Open space	People
	E	Stairs	Find starting point of stairs

variety of tasks performed in controlled laboratory settings, for instance, learning to navigate using hypertext and multimedia data [14–16] and learning to use complex maps during hypermedia navigation [17]. Despite EEG's ability to capture the cognitive load when performing a task, its usability in commercial assistive devices is still in its infancy. For this reason, we collected a wide range of physiological signals, such as skin conductance by means of a wearable bracelet. Based on findings in the literature, skin conductance may predict the performance in a task under stressful conditions [18–20]; confirming such statement in "out-of-laboratory" conditions brings great advantages in the design of assistive devices.

We contribute to the existing literature by conducting navigation experiments exclusively "in the wild," where the VIP participants navigated in predefined indoor and outdoor routes previously unfamiliar to the them. These routes included a large variety of obstacles and different urban environments (see Table 2). A machine learning framework was designed based on random forest classifiers to predict the cognitive load of the participants for each time instance inferring on physiological features extracted from the skin conductance signals. The aim of this study is twofold; first, we exploit possible effects that the various urban indoor and outdoor environments may induce on people in relation to their degree of sight loss and, second, to pinpoint easily accessible biomarkers that robustly predict the cognitive load of VIP when navigating in unfamiliar sites in the wild.

In line with the current literature [18–20], the emerging cross-validated results suggest that physiological features related to skin conductance are accurately and robustly predicting the amount of cognitive load in real time. Taking into consideration these findings, the design of assistive aids adapts in real time to the requirements and personal needs of the user.

2. Data Collection

2.1. Participants. A total of ten healthy visually impaired adults with different degrees of sight loss participated in the two mobility studies (6 females; average age = 41 yrs, range = 22–53 yrs). To help make them feel comfortable and safe, they were encouraged to walk as usual using their white canes if they wished so and were accompanied by their familiar O&M instructor. Participants were instructed to avoid smoking normal or e-cigarettes and consuming caffeine or sugar (e.g., coffee, coke, and chocolate) approximately one hour prior to the walk. Recruitment was based on volunteering and all VIP were capable of giving free and informed consent. The study was approved by the National Bioethics Committee of Iceland. All data was anonymized before analysis. Seven of the participants walked both the outdoor and indoor routes, one took part only in the outdoor study, and two completed only the indoor task (see Table 1).

2.2. Indoor and Outdoor Routes. The indoor experiment was conducted inside a building of the University of Iceland in Reykjavik. With the assistance of VIP caretakers and O&M instructors, we planned a route to take the VIP through circumstances where different levels of stress were likely to

occur (i.e., of varying complexity and difficulty). Participants walked the charted route three times for training purposes. The route comprised five distinct environments representable of a variety of indoor mobility challenges (see Table 2). Indicatively, participants had to enter through automated doors, use an elevator, move across a busy open space, walk down a large spiral staircase, and walk through other obstacles. The route was approximately 200 meters in length and took on average 5 minutes to walk (range = 4–8 minutes).

The outdoor route was charted in the city center of Reykjavik in Iceland. It comprised eight distinct scenes defined so as to cluster environmental and situational factors expected to elicit similar affective reactions. For example, participants had to walk on a busy shopping street, stroll through an urban park, cross a major junction, and pass through narrow sidewalks (see Table 2). The route was approximately 1 km long and took on average 13 min 44 s to walk (range = 9–19 min).

2.3. Multimodal Biosignals. EEG signals were recorded using the Emotiv EPOC+ (http://emotiv.com/epoc/), a mobile headset with 16 dry electrodes registering over the 10-20 system locations AF3, F7, F3, FC5, T7, P3 (CMS), P7, O1, O2, P8, P4 (DRL), T8, FC6, F4, F8, and FC4 (sampling rate $f_s = 128$ Hz). Given the practical constraints involved in monitoring brain electrical activity in the wild, EPOC+ was chosen because it provides a good compromise between performance (i.e., number of channels and scientific validity of the acquired EEG signals) and usability (i.e., portability, preparation time, and user comfort) with respect to other commercial wireless EEG systems [21–24].

Along with the Emotiv headset, participants were asked to wear the Empatica E4 wristband (https://www.empatica.com/e4-wristband) [25]. E4 measures EDA as skin conductance through 2 ventral (inner) wrist electrodes ($f_s = 4$ Hz) and BVP through a dorsal (outer) wrist photoplethysmography (PPG) sensor ($f_s = 64$ Hz). E4 further reports HR, extracted on board from BVP interbeat intervals. The wristband also includes an infrared thermopile sensor and a 3-axis accelerometer. E4 is currently the only commercial multisensor device developed based on extended scientific research in the areas of psychophysiology and affective computing. Additionally, it has a cable-free, watch-like design, which makes it easier and more aesthetically pleasing to wear and thus better fitted to use in the wild compared to other wearable biosignal devices. Participants were asked to wear the wristband on the nondominant hand to minimize motion artifacts related to handling the white cane [26].

2.4. General Procedure. Participants walked the outdoor route twice and the indoor route three times for training purposes. In both studies directions were only provided during the first walk to help the VIP familiarize with the route. They were instructed to avoid unnecessary head movements and hand gestures as well as talking to their O&M instructor unless there was an emergency. Video and audio were registered by means of a smartphone camera to facilitate data annotation (observing behaviors across the different environments and situations) and synchronization (start/end of walk, environments, and obstacles). In the outdoor study, GPS coordinates were additionally logged using a Garmin GPSMAP-64s unit at a rate of 1 registration per second. Upon completing the last walk, participants were asked to describe stressful moments they experienced along the route.

3. Feature Extraction

3.1. EEG. The EEG data was first time-domain interpolated using the Fast Fourier Transform (FFT) to account for missing samples due to connectivity issues. Subsequently, all signals were baseline-normalized by subtracting for each participant and for each channel the mean of resting state registrations. These were obtained during a series of laboratory studies with the same participants [27, 28].

Based on findings in the neuroscientific literature we extracted a series of features descriptive of the cognitive and the physiological state of the participants in each time instance. The brain activity is characterized by rhythmic patterns across distinct frequency bands, the definition of which can vary somewhat among studies. Here we analyzed EEG in six bands, namely, delta (0.5–4 Hz), theta (4–7 Hz), alpha 1 (7–10 Hz), alpha 2 (10–13 Hz), beta (13–30 Hz), and gamma (30–60 Hz). Beta activity is associated with psychological and physical stress, whereas theta and alpha 1 (i.e., lower alpha) frequencies reflect response inhibition and attentional demands such as phasic alertness [29]. Alpha 2 (i.e., higher alpha) is related to task performance in terms of speed, relevance, and difficulty [30]. Gamma waves are involved in more complex cognitive functions such as multimodal processing or object representation [31]. Features related to signal power and complexity were extracted using the PyEEG open source Python module [32]. For each of the 14 EEG channels, we computed the *Relative Intensity Ratio* as an indicator of relative spectral power in each of the six frequency bands [33].

Having extracted the power band features from the EEG signals, we estimated the event-related (de)synchronization (ERD/ERS) index, a well-established measure of band power change in EEG originally proposed by Pfurtscheller and Aranibar [34]. It is defined as

$$\frac{\text{ERD}}{\text{ERS}}(\%) = \frac{\text{baseline IBP} - \text{test IBP}}{\text{baseline IBP}} * 100, \qquad (1)$$

where IBP stands for interval band power. The baseline IBP refers to a prestimulus time period without any task demands, in our case the resting state, whereas the activation interval (test IBP) refers to the time period while working on the experimental task. We slightly modified the estimation of the ERD/ERS index, defining test IBP as the time interval of one second of our recorded data. In this way, we result with one time point of ERD/ERS per second, where every time point expresses the synchronization or desynchronization according to the same baseline.

3.2. EDA. The skin conductance data was decomposed into two continuous components, namely, phasic and tonic component [35]. This decomposition and subsequent extraction

of tonic and phasic electrodermal activity (EDA) features were performed using the Ledalab toolbox (http://www.ledalab.de/). Overall, we extracted six features: mean tonic EDA (TM) and the number of "spontaneous" SCRs (i.e., phasic changes not traceable to specific stimulation), which are known to be particularly suitable for longitudinal monitoring of emotional stress-elicited EDA (i.e., tonic arousal); sum of amplitudes of registered SCRs (AS) and average, maximum, and cumulative phasic EDA (PM), which provide varying indicators of instantaneous phasic arousal [26].

3.3. BVP and HR. The photoplethysmography sensor of the E4 device measures the blood volume pulse (BVP) from which it derives on board the heart rate (HR). We min-max normalized both data streams to account for interindividual differences [36].

4. Linear Mixed Model Analysis

4.1. Method. To examine differences in mental activity between outdoor and indoor scenes of varying complexity and obstacles in relation to the amount of vision loss, a linear mixed model analysis was conducted for the alpha 1 (lower alpha) and alpha 2 (upper alpha) bands in each of the two routes (outdoor, indoor). Linear mixed models perform a regression-like analysis while controlling for random variance caused by differences in factors such as participant and electrode [37, 38]. We chose to focus on the alpha bands only because it has been repeatedly observed that brain activity at those frequencies is associated with cognitive load in a variety of task demands: specifically, alpha activity has been shown to fall in magnitude (i.e., alpha ERD increases) with higher task difficulty (see [39] for a review).

Fixed factors examined in the analysis included type of scene (Table 2) and category of vision impairment (Table 1). For the latter, two broader categories of vision loss were considered to better fit the linear models to the data: almost blind (categories VI-5 and VI-4) and severely impaired (categories VI-3 and VI-2). Random intercepts for each participant and electrode position were added. Type III Wald F-tests were used to test the significance of the fixed factors and their interaction [40]. Pairwise comparisons of group means were carried out with t-tests, using Bonferroni-adjusted p values where appropriate [41]. Before averaging across conditions, a logarithmic transformation of single-condition ERD/ERS values was applied to improve their distributional characteristics.

4.2. Results. The across-participants average ERD/ERS values for each environment and for each category of vision impairment are shown in Figure 1. For each subplot, mean values for outdoor scenes are depicted in the left panel, whereas those for indoor environments are drawn in the right panel. Type III Wald F-test results from the four (two bands × two routes) linear mixed models are reported in Table 3. Vision alone was only a significant predictor of upper alpha ERD/ERS in the outdoor route, although the interaction of vision and scene was significantly influential for both lower and upper alpha ERD/ERS in the outdoor route. The scene alone had a significant effect on both bands and in both outdoor and indoor scenarios.

Post hoc paired samples t-tests showed that ERD/ERS in the lower alpha band was significantly higher for almost blind than for severely impaired individuals for the outdoor environments B [small street; $t(13.53) = 2.43$, $p = 0.030$], E [open space; $t(10.87) = 2.75$, $p = 0.019$], and G [crossing small street without traffic lights; $t(14.78) = 2.86$, $p = 0.012$]. Similar trends were found for outdoor ERD/ERS in the upper alpha band [B: $t(20.51) = 4.13$, $p = 0.001$; E: $t(15.85) = 2.14$, $p = 0.048$; G: $t(25.92) = 4.20$, $p < 0.001$]. In addition, upper alpha ERD/ERS was found to be significantly higher for almost blind than for severely impaired participants for the outdoor environments A [shopping street; $t(13.52) = 2.35$, $p = 0.035$] and H [construction alley; $t(24.07) = 2.47$, $p = 0.021$]. For the indoor environments, lower alpha EDR/ERS was only significantly higher for almost blind than for severely impaired individuals when walking up and down stairs [scene E; $t(16.83) = 2.23$, $p = 0.040$].

When averaging across the two VI groups, lower alpha ERD/ERS was significantly higher when crossing a main traffic junction than when passing through the shopping street [$t(846.34) = 3.23$, $p = 0.036$], small street [$t(846.47) = 3.27$, $p = 0.032$], and small street crossing scenes [$t(847.22) = 3.22$, $p = 0.038$]. ERD/ERS in the lower alpha band was higher when passing through the shopping street than the small street and higher for the latter than when crossing a small street, but these differences were not found to be significant. Similar trends were obtained for upper alpha ERD/ERS in the outdoor model [$3.68 < t(833.64\text{--}834.56) < 4.44$, $p < 0.007$], while significantly higher upper alpha ERD/ERS was also observed for the urban park scene compared to the small street environment [$t(833.74) = 3.31$, $p = 0.027$]. For the indoor route, ERD/ERS in the lower alpha band was significantly higher when using automated moving doors and when taking the elevator than when walking along a narrow corridor [$t(922.63) = 3.27$, $p = 0.011$ and $t(922.43) = 3.48$, $p = 0.005$, resp.], navigating through an open space [$t(922.57) = 5.06$, $p < 0.001$ and $t(920.34) = 5.32$, $p < 0.001$, resp.], and using the stairs [$t(914.80) = 5.15$, $p < 0.001$ and $t(912.93) = 5.47$, $p < 0.001$, resp.]. Lower alpha ERD/ERS was higher for the elevator than for the door scene, but not significantly so. It was higher for the corridor than for the stairs environments and higher for the latter than for the open space scene, but these differences were also not found to be significant. Upper alpha ERD/ERS was also significantly higher when using automated moving doors and when taking the elevator than in the other indoor environments [$3.19 < t(887.33\text{--}893.52) < 7.13$, $p < 0.015$ and $3.22 < t(886.08\text{--}893.01) < 7.27$, $p < 0.014$, respectively], while trends similar to the outdoor model were observed for the remaining indoor scene contrasts.

Overall, outdoor and indoor environments that were more dynamic with respect to complexity and unexpected obstacles, such as crossing a major road, strolling through an open urban space, walking through a narrow alley with coffee tables and advertisement boards, using an elevator, and going through automatic doors, resulted in substantially higher ERD values (i.e., lower relative power) across the two alpha

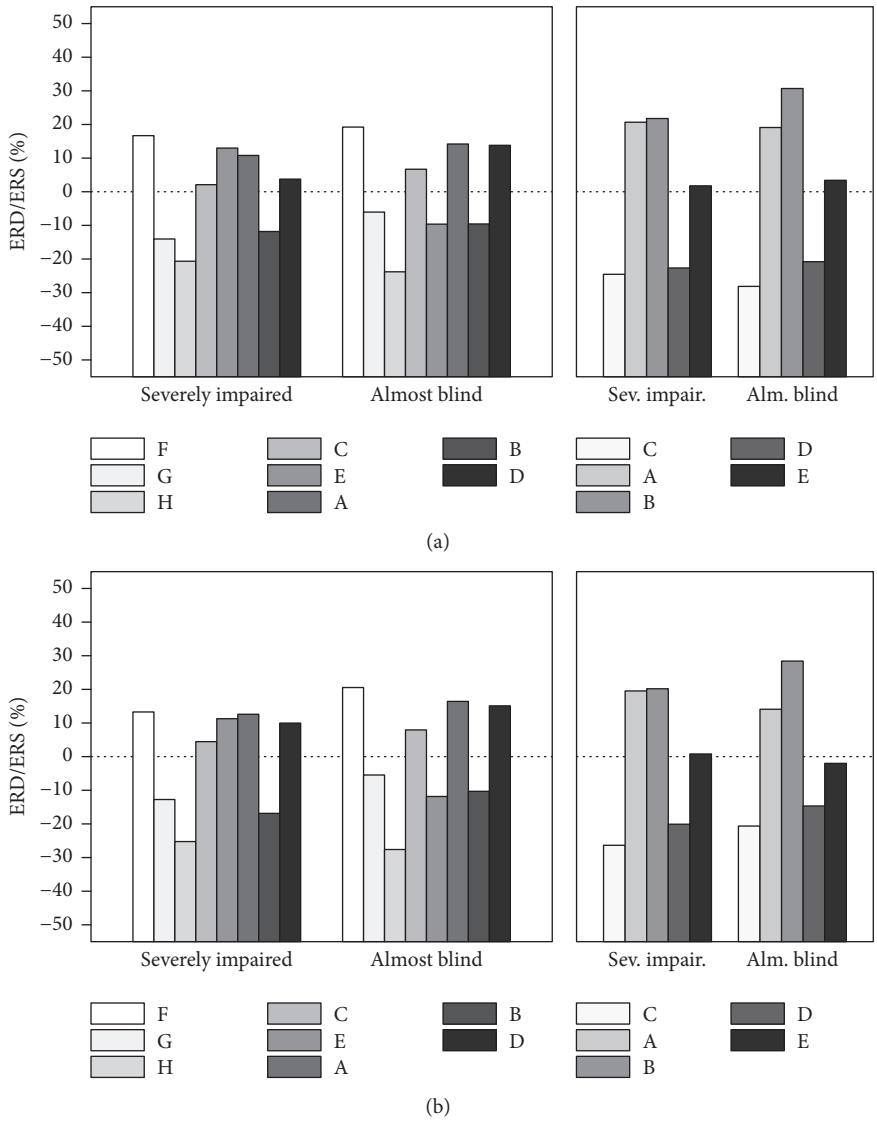


Figure 1: Average ERD/ERS values for each scene (Table 2) and for each category of vision impairment (severely impaired: visual acuity less than 10% but greater than 2%; almost blind: visual acuity less than 2%). (a) Lower alpha band ERD/ERS. (b) Upper alpha band ERD/ERS. In each subplot, outdoor scenes are depicted in the left panel and indoor environments are drawn in the right panel.

Table 3: Linear model type III Wald F-tests for ERD/ERS in the alpha-1 and alpha-2 bands, in the outdoor and indoor routes.

	df	F	p	df	F	p
		alpha-1, outdoor			*alpha-1, indoor*	
Intercept (I)	1, 7.91	620.84	<0.001	1, 8.21	1074.04	<0.001
Vision Impair (VI)	1, 6.02	2.61	0.157	1, 7.58	2.62	0.147
Scene (S)	7, 846.56	2.83	0.006	4, 916.92	14.58	<0.001
VI × S	7, 846.36	5.15	<0.001	4, 915.99	1.09	0.360
		alpha-2, outdoor			*alpha-2, indoor*	
I	1, 7.97	1027.09	<0.001	1, 9.74	903.41	<0.001
VI	1, 6.05	7.38	0.035	1, 7.37	1.23	0.302
S	7, 833.86	5.39	<0.001	4, 888.74	20.92	<0.001
VI × S	7, 833.90	7.30	<0.001	4, 887.90	1.16	0.325

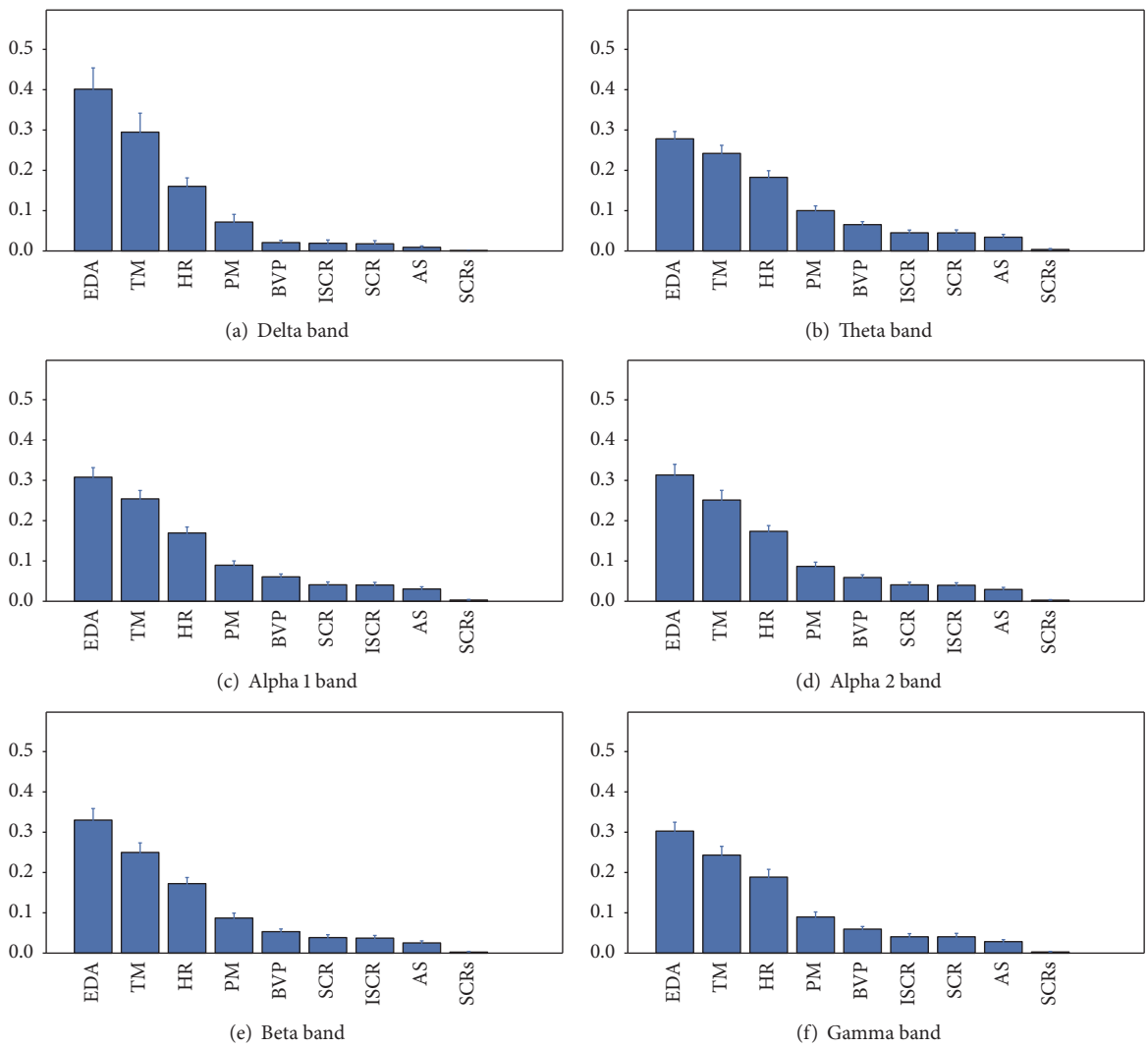


Figure 2: Feature importances as emerging from the Gini criterion of each RF model. The error bars refer to the Gini value over 5-fold. Note that EDA, mean tonic EDA (TM), and heart rate (HR) emerge as the most predictive biomarkers for the prediction of the cognitive load index all the frequency bands showing the stability of the approach.

bands, which implies increased task difficulty. These cognitive load "hotspots" are in full agreement with the scenes reported as stressful by the participants themselves at the end of the study.

5. Automatic Prediction of Cognitive Load

5.1. Classification Experiments. To automatically identify the cognitive load of urban indoor and outdoor spaces experienced by VIP while walking through it based only on their biosignals, we postulated the study as a supervised classification process. A widely used ensemble learning method for classification was employed, namely, Random Forest (RF) classifier [42], selected due to its ability to deal with possibly correlated predictor variables and because it provides a straightforward assessment of the variable importances.

The ERD/ERS index of cognitive load was averaged over all electrodes per frequency band per second. The resulting averaged index was binned in three chunks, namely, "Low," "Medium," and "High" load. We trained a RF model to predict the aforementioned labels of cognitive load index per each band, inferring on the features extracted from the skin conductance and blood volume pulse sensor. The adjustment of the two most important parameters of RF was performed by means of grid search parameter estimation with 5-fold cross-validation. We exploited the effect of the number of estimators $[150, 300, 600]$ and the effect of the maximum number of features $[.5, 1, 2] * \sqrt{\text{Number Of Features}}$. Overall, the optimum number of estimators was 300 and the maximum number of features was set equal to the total number of features for each experiment.

5.2. Results. Table 4 reports the classification results in terms of AUROC weighted metric. Hereafter, we will refer to AUROC weighted metric with the term "accuracy." For each frequency band, the average accuracy over 5-fold is reported,

Table 4: Classification AUROC weighted metric for all the environments across the various experiments. The reported numbers refer to the mean AUROC over all folds in percentile and in parenthesis the standard deviation is reported.

Frequency band	AUROC weighted average (SD)
Delta	0.97 (0.00)
Theta	0.83 (0.00)
Alpha 1	0.85 (0.00)
Alpha 2	0.85 (0.00)
Beta	0.86 (0.00)
Gamma	0.84 (0.00)

along with the respective standard deviation. We note that for all frequency bands the performance of the models is quite accurate and robust.

As mentioned, the ERD/ERS index employed for the definition of the classes was averaged over all electrodes; in literature, there are many studies associating specific electrodes to brain functions, for instance, C_z to memory recall tasks; however, the Emotiv EPOC+ used for the experiments does not provide a full coverage of the cranial surface so as to focus on specific electrodes. Following the exact same scheme for the classification of the cognitive load states ("Low," "Medium," and "High") from the separate electrodes per band we obtained accuracy values identical to the averaged results per band.

Figure 2 depicts the most predictive features of EDA and heart rate of the cognitive load. Note that the order of importance and the relative amplitude of the "Gini" importance value are comparable for all the frequency bands showing the stability of the approach. These findings are in line with the studies in the literature, where the skin resistance is stated to be an important indicator of the cognitive load [18–20].

6. Conclusions

This paper presents a framework for real-time automatic assessment of cognitive load when visually impaired people move and navigate in unfamiliar outdoor and indoor environments. The objective is to demonstrate the feasibility of real-time tracking of mentally demanding tasks which can be used as on the fly feedback to assistive devices. Mobility aids for visually impaired people should be capable of implicitly adapting not only to changing environments but also to shifts in the cognitive load of the user in relation to different environmental and situational factors.

The proposed framework is based on multimodal fusion of brain and peripheral biosignal features. Using stress-related features of the EDA signal and an EEG index of cognitive load based on event-related (de)synchronization in the alpha band (ERD/ERS), we identified the most important cognitively demanding "hotspots" for the generic VIP population and for the specific categories of sight loss, pointing out the particular needs/difficulties faced by each VIP category. The high prediction rates in the multimodal classification experiments (83–97% AUROC Weighted, Table 4) are very encouraging of the proposed approach. Even if the chosen urban and building sites did not represent all possible different outdoor and indoor environments and situations in terms of complexity and difficulty, the charted routes were designed so as to combine most of the mobility challenges faced by VIP.

Despite being promising, reported findings should be considered with caution due to the limited number of participants, which did not allow for an in-depth analysis of specific stressors in each category of vision impairment. A larger group study would need to be carried out to confirm and quantify the trends obtained here. Furthermore, the well-established Emotiv EPOC+ EEG headset has certain limitations with respect to the quality of the recorded signal during experiments involving physical activity "in the wild" such as those presented here. Future steps of the present study include refining the predictive model through exploring novel multimodal biosignal features for cognitive load assessment and comparing different classifiers. Such findings hopefully pave the way to emotionally intelligent mobile technologies that take the concept of navigation one step further, accounting not only for the shortest path but also for the most effortless, least stressful, and safest one.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors wish to thank the administration and O&M instructors at the National Institute for the Blind, Visually Impaired, and Deaf in Iceland for their valuable input and generous assistance and the visually impaired individuals who took part in the study for their time and patience. The research leading to these results has received funding from the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement no. 643636 "Sound of Vision." Charalampos Saitis acknowledges the Alexander von Humboldt Foundation for support through a Humboldt Research Fellowship.

References

[1] World Health Oganization (WHO) Media Centre, "Visual Impairment and Blindness Fact Sheet no. 282, 2014," http://www.who.int/mediacentre/factsheets/fs282/en/Available [Accessed February 13, 2014].

[2] R. William Wiener and E. Siffermann, "assessment," in *Foundations of orientation and mobility*, pp. 272–275, 1997, assessment.

[3] A. Virtanen and S. Koskinen, "Noppa: navigation and guidance system for the visually impaired," in *Proceedings of the 11th World Congress and Exhibition on ITS*, 2004.

[4] R. G. Golledge, J. M. Loomis, R. L. Klatzky, A. Flury, and X. L. Yang, "Designing a personal guidance system to aid navigation without sight; progress on the gis component," *International Journal of Geographical Information Science*, vol. 5, no. 4, pp. 373–395, 1991.

[5] A. Rouben and L. Terveen, *Speech and Non-Speech Audio: Navigational Information and Cognitive Load*, Georgia Institute of Technology, 2007.

[6] C. Jacquet, Y. Bellik, and Y. Bourda, "Electronic locomotion aids for the blind: Towards more assistive systems," *Studies in Computational Intelligence*, vol. 19, pp. 133–163, 2006.

[7] Bay Advanced Technologies Ltd., "The Bat 'k' Sonar-Cane," http://www.eastin.eu/en/searches/products/detail/database-vli-bank/id-E21029.

[8] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: a survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 1, pp. 25–35, 2010.

[9] R. Velázquez, "Wearable assistive devices for the blind," *Lecture Notes in Electrical Engineering*, vol. 75, pp. 331–349, 2010.

[10] N. A. Giudice and G. E. Legge, "Blind navigation and the role of technology," in *The Engineering Handbook of Smart Technology for Aging, Disability, and dependence*, A. Helal, M. Mokhtari, and B. Abdulrazak, Eds., pp. 479–500, John Willey & Sons, 2008.

[11] C. Saitis and K. Kalimeri, "Identifying urban mobility challenges for the visually impaired with mobile monitoring of multimodal biosignals," in *Universal Access in Human-Computer Interaction. Users and Context Diversity*, M. Antona and C. Stephanidis, Eds., pp. 616–627, Springer, 2016.

[12] K. Kalimeri and C. Saitis, "Exploring multimodal biosignal features for stress detection during indoor mobility," in *Proceedings of the 18th ACM International Conference on Multimodal Interaction, ICMI 2016*, pp. 53–60, Japan, November 2016.

[13] J. Sweller, "Cognitive load during problem solving: effects on learning," *Cognitive Science*, vol. 12, no. 2, pp. 257–285, 1988.

[14] D. O. Bos et al., "Eeg-based emotion recognition," *The Influence of Visual and Auditory Stimuli*, vol. 56, 1, no. 3, p. 17, 2006.

[15] C. Berka, D. J. Levendowski, M. N. Lumicao et al., "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks," *Aviation, Space, and Environmental Medicine*, vol. 78, no. 5, pp. B231–B244, 2007.

[16] R. M. Nilsson and R. E. Mayer, "The effects of graphic organizers giving cues to the structure of a hypertext document on users' navigation strategies and performance," *International Journal of Human-Computer Studies*, vol. 57, no. 1, pp. 1–26, 2002.

[17] B. M. Scott and N. H. Schwartz, "Navigational spatial displays: the role of metacognition as cognitive load," *Learning and Instruction*, vol. 17, no. 1, pp. 89–105, 2007.

[18] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable eda device," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 410–417, 2010.

[19] C. Mundell, J. Pablo Vielma, and T. Zaman, *Predicting Performance under Stressful Conditions Using Galvanic Skin Response*, 2016.

[20] Y. Shi, N. Ruiz, R. Taib, E. Choi, and F. Chen, "Galvanic skin response (GSR) as an index of cognitive load," in *Proceedings of the 25th SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, pp. 2651–2656, ACM, San Jose, Calif, USA, May 2007.

[21] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur, "Validation of the emotiv EPOC® EEG gaming systemfor measuring research quality auditory ERPs," *PeerJ*, vol. 1, no. 1, article e38, 2013.

[22] W. David Hairston, K. W. Whitaker, A. J. Ries et al., "Usability of four commercially-oriented EEG systems," *Journal of Neural Engineering*, vol. 11, no. 4, p. 046018, 2014.

[23] S. Debener, F. Minow, R. Emkes, K. Gandras, and M. de Vos, "How about taking a low-cost, small, and wireless EEG for a walk?" *Psychophysiology*, vol. 49, no. 11, pp. 1617–1621, 2012.

[24] J. I. Ekandem, T. A. Davis, I. Alvarez, M. T. James, and J. E. Gilbert, "Evaluating the ergonomics of BCI devices for research and experimentation," *Ergonomics*, vol. 55, no. 5, pp. 592–598, 2012.

[25] M. Garbarino, M. Lai, S. Tognetti, R. Picard, and D. Bender, "Empatica E3 - A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition," in *Proceedings of the 4th International Conference on Wireless Mobile Communication and Healthcare - "Transforming healthcare through innovations in mobile and wireless technologies"*, Athens, Greece, November 2014.

[26] W. Boucsein, *Electrodermal Activity*, Springer, New York, NY, USA, 2nd edition, 2012.

[27] S. Spagnol, O. I. Johannesson, A. Kristjansson et al., "Model-based obstacle sonification for the navigation of visually impaired persons," in *Proceedings of the 19th International Conference on Digital Audio Effects, DAFx*, pp. 309–316, Brno, Czech Republic, September 2016.

[28] S. Spagnol, C. Saitis, K. Kalimeri, I. Omar, and R. Unnthorsson, "Sonificazione di ostacoli come ausilio alla deambulazione di non vedenti (Obstacle sonification as navigation aid for the visually impaired)," in *Proceedings of XXI Colloquio di Informatica Musicale (Colloquium on Music Informatics)*, Cagliari, Italy, 2016.

[29] W. J. Ray and H. W. Cole, "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes," *Science*, vol. 228, no. 4700, pp. 750–752, 1985.

[30] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Research Reviews*, vol. 29, no. 2-3, pp. 169–195, 1999.

[31] A. Keil, M. M. Müller, W. J. Ray, T. Gruber, and T. Elbert, "Human gamma band activity and perception of a gestalt," *The Journal of Neuroscience*, vol. 19, no. 16, pp. 7152–7161, 1999.

[32] F. S. Bao, X. Liu, and C. Zhang, "PyEEG: an open source python module for EEG/MEG feature extraction," *Computational Intelligence and Neuroscience*, vol. 2011, Article ID 406391, 2011.

[33] R. Quian Quiroga, S. Blanco, O. A. Rosso, H. Garcia, and A. Rabinowicz, "Searching for hidden information with gabor transform in generalized tonic-clonic seizures," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 4, pp. 434–439, 1997.

[34] G. Pfurtscheller and A. Aranibar, "Event-related cortical desynchronization detected by power measurements of scalp EEG," *Electroencephalography and Clinical Neurophysiology*, vol. 42, no. 6, pp. 817–826, 1977.

[35] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of Neuroscience Methods*, vol. 190, no. 1, pp. 80–91, 2010.

[36] J. T. Cacioppo and L. G. Tassinary, "Inferring psychological significance from physiological signals," *American Psychologist*, vol. 45, no. 1, pp. 16–28, 1990.

[37] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, no. 4, pp. 963–974, 1982.

[38] D. Bates, M. Mächler, B. M. Bolker, and S. C. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software* , vol. 67, no. 1, 2015.

[39] P. Antonenko, F. Paas, R. Grabner, and T. van Gog, "Using electroencephalography to measure cognitive load," *Educational Psychology Review*, vol. 22, no. 4, pp. 425–438, 2010.

[40] J. Fox and W. Sanford, *An R Companion to Applied Regression*, Sage, 2nd edition, 2011.

[41] R. V. Lenth, "Least-squares means: The R package lsmeans," *Journal of Statistical Software* , vol. 69, 2016.

[42] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

Hindawi
Wireless Communications and Mobile Computing
Volume 2017, Article ID 2986423, 16 pages
https://doi.org/10.1155/2017/2986423

Research Article

Wireless Brain-Robot Interface: User Perception and Performance Assessment of Spinal Cord Injury Patients

Alkinoos Athanasiou,[1,2] George Arfaras,[1] Niki Pandria,[1] Ioannis Xygonakis,[1] Nicolas Foroglou,[2] Alexander Astaras,[1,3] and Panagiotis D. Bamidis[1]

[1] *Biomedical Electronics Robotics & Devices (BERD) Group, Lab of Medical Physics, School of Medicine, Faculty of Health Sciences, Aristotle University of Thessaloniki (AUTH), 54124 Thessaloniki, Greece*
[2] *1st Department of Neurosurgery, "AHEPA" University General Hospital, Aristotle University of Thessaloniki (AUTH), 54636 Thessaloniki, Greece*
[3] *Robotics Laboratory, Computer Science Department, American College of Thessaloniki (ACT), 55535 Thessaloniki, Greece*

Correspondence should be addressed to Alkinoos Athanasiou; athalkinoos@auth.gr

Received 25 August 2017; Accepted 10 December 2017; Published 31 December 2017

Academic Editor: Kyriaki Kalimeri

Patients suffering from life-changing disability due to Spinal Cord Injury (SCI) increasingly benefit from assistive robotics technology. The field of brain-computer interfaces (BCIs) has started to develop mature assistive applications for those patients. Nonetheless, noninvasive BCIs still lack accurate control of external devices along several degrees of freedom (DoFs). Unobtrusiveness, portability, and simplicity should not be sacrificed in favor of complex performance and user acceptance should be a key aim among future technological directions. In our study 10 subjects with SCI (one complete) and 10 healthy controls were recruited. In a single session they operated two anthropomorphic 8-DoF robotic arms via wireless commercial BCI, using kinesthetic motor imagery to perform 32 different upper extremity movements. Training skill and BCI control performance were analyzed with regard to demographics, neurological condition, independence, imagery capacity, psychometric evaluation, and user perception. Healthy controls, SCI subgroup with positive neurological outcome, and SCI subgroup with cervical injuries performed better in BCI control. User perception of the robot did not differ between SCI and healthy groups. SCI subgroup with negative outcome rated Anthropomorphism higher. Multi-DoF robotics control is possible by patients through commercial wireless BCI. Multiple sessions and tailored BCI algorithms are needed to improve performance.

1. Introduction

Spinal Cord Injury (SCI) is a potentially life-changing condition, causing permanent disability, compromising the victim's physical and psychological well-being and impacting their close environment as well. Brain-computer interfaces (BCIs) are rapidly developing into a field-changing technology for those patients, not only replacing motor functions [1, 2] but even promising to alter the long-term outcome of the condition [3, 4]. Wireless invasive (implantable) neural recording is also an important development [5, 6], especially considering that SCI has become the research target of several assistive technologies [7] including functional electrical stimulation [8] and robotics for neurorehabilitation [9].

On the other hand, noninvasive BCI technology is dominated by the EEG modality, which has great temporal but low spatial resolution. EEG can detect rapid dynamics of the brain but lacks source estimation. Consequently, a major limitation of noninvasive BCI is the low signal to noise ratio (SNR) which can be partially attributed to volume conduction effect [10, 11]. Since the EEG signal measured at the scalp is the superposition of all electrical signals, including those generated by the cortex, discriminating brain activity from artifacts and noise can be technically difficult.

In addition to these neurophysiological limitations, a major drawback of traditional noninvasive BCIs is inconvenience. The subject's movement is typically constrained by wires connecting the EEG electrode cap to the data

acquisition unit, while the entire setup of wet electrodes entails a time-consuming procedure. From the hardware perspective, the main challenges of noninvasive BCI technology are portability and ease-of-use, rendering the wireless dry-electrode cap a promising solution. Lessons could be derived from the—otherwise immobile—MEG modality that boasts contactless sensors and reduced preparation time [13]. In terms of performance, the main challenge for noninvasive BCI technology is to accurately control external devices (multijoint robotic arms, drones, wheelchairs, etc.) that can move along several degrees of freedom (DoFs), a task which is necessary for performing complex operations in the physical world. Such a feat requires highly accurate decoding algorithms, able to discriminate multiple classes of Motor Imagery (MI) under the limitations of low SNR and spatial resolution. It has been demonstrated that the basic operation of a robotic arm which fits the aforementioned description is possible using noninvasive BCI [14].

Continuous advancements in electronics, from solid-state transistors in the 1940s, integration of sensors and powerful digital microprocessors in the 1970s, and other developments in the following decades, such as switching frequency, increased computing power, and programming flexibility, decreased fabrication cost and power consumption and steadily led to lighter and more agile, responsive, and computationally complex robots. The demand in the fields of Medical Rehabilitation and Assistive Technologies (AT) leads to the first combined medical applications for paralysis and stroke patients: robotic sleeves for assisted living, powered orthotics, and even some initial attempts at robotic manipulation arms [15, 16]. This nascent field of Rehabilitation Robotics started taking shape in the 1970s and 1980s, gradually emphasizing novel robotic manipulator designs [17, 18] and attempts to adapt existing industrial robotic manipulators [19].

Unobtrusiveness, portability, and mobility are clear advantages for AT and Medical Robotics. Even though affordable, power-autonomous, full-body robotic exoskeletons are yet to materialize, disabled patients now have a variety of product options, some commercially available: manipulator arms mounted on robotic wheelchairs [20, 21], dexterous robotic prosthetics which can acquire their control signals directly from the patient's nervous system [1, 22], and adaptations of aforementioned BCI technology specifically for robotic rehabilitation [23–25]. Unfortunately, research in optimal interfaces for Human-Robot Interaction (HRI) is often overlooked in medical robotics development [12], with a potentially significant impact on user acceptance and the validation success rate for new technologies [26]. User perception and overall satisfaction with a robot's technological interface has equal importance to hardware/software design and quality standards, something that holds particularly true in robotic-assisted rehabilitation. The need for patient's immersive experience and willingness to collaborate with the robot, physiotherapist, and supervising physician should meaningfully alter the criteria of success of a robot [27]. For instance, acceptance of an external machine as part of one's own body schema can significantly impact the rehabilitation process and should also be taken into account [15, 28].

In our previous work we have presented our progress towards developing 8-DoF anthropomorphic robotic arms, controlled by wireless off-the-shelf BCI, for AT and rehabilitation applications [12]. We have accounted for development of the robotic arms and electronics, for implementation of the BCI control module, and we have presented pilot experimental applications of the Brain-Robot Interface (BRI) on healthy and disabled individuals [29, 30].

In the remainder of this paper we present an elaborate user-assessment study of our wireless BRI by subjects with SCI and healthy individuals [31]. We focus not only on performance assessment during multiple movements BCI control, but on user perception of the assistive technology as well, analyzed with regard to their neurological condition, independence, imagery capacity, and psychometric evaluation. In Materials and Methods, we briefly present the BRI and we explain the experimental setup and statistical analysis of our collected data, which we then present in Results. In Discussion we attempt to interpret our results, comment on the strengths, and underline the limitations of our approach. We also discuss planned technological development in the direction of robotics, BCI module, signal analysis, and further experiments, as well as the challenges we are yet to meet.

2. Materials and Methods

2.1. The Brain-Robotic Arms Interface. As both the interface's and robotics' development have been thoroughly described before [12, 32], we will briefly explain the system's technical characteristics here, emphasizing wireless capabilities. The Emotiv EPOC is an easily applied wireless 14-saline felt-electrode EEG system (Emotiv, USA), capable of detecting brain activity at a bandwidth of 0.2–43 Hz, employing digital notch filters at 50 Hz and 60 Hz and built-in digital 5th order Sinc filter. Meanwhile, due to being commercial class, it costs significantly less and also is less obtrusive, simple to use, and portable compared to traditional medical EEG devices. Its technical characteristics are also on par with the designated use case, so our team used this device for the development of the BCI modality. The wireless chipset transmits at a proprietary ultra-low energy (ULE) 2.4 Ghz Bluetooth Smart to a USB dongle connected to a dedicated computing laptop. Internal sampling rate of the electrodes is at 2 kHz (sequential sampling and single analog to digital conversion) with 16-bit resolution and signal sampling rate of 128 Hz transmitted. Using the Cognitiv suite a resting state and up to four different mental commands are classified (using proprietary BCI algorithm) and then are appended to key button input through the Emokey software. The input then accesses a control MATLAB script corresponding to movement commands for the robotic arms (Figure 1). Movement coordinates are subsequently transmitted to the robotic arms system through a serial port with a Baud Rate of 9600 bps.

The "Mercury 2.0" robotic arm is a stand-alone electromechanical manipulator system developed by our team, capable of replicating most movements of a physical human arm. The current version of the robot is capable of movement along 8 DoFs [32]. Eight motors are attached in total on

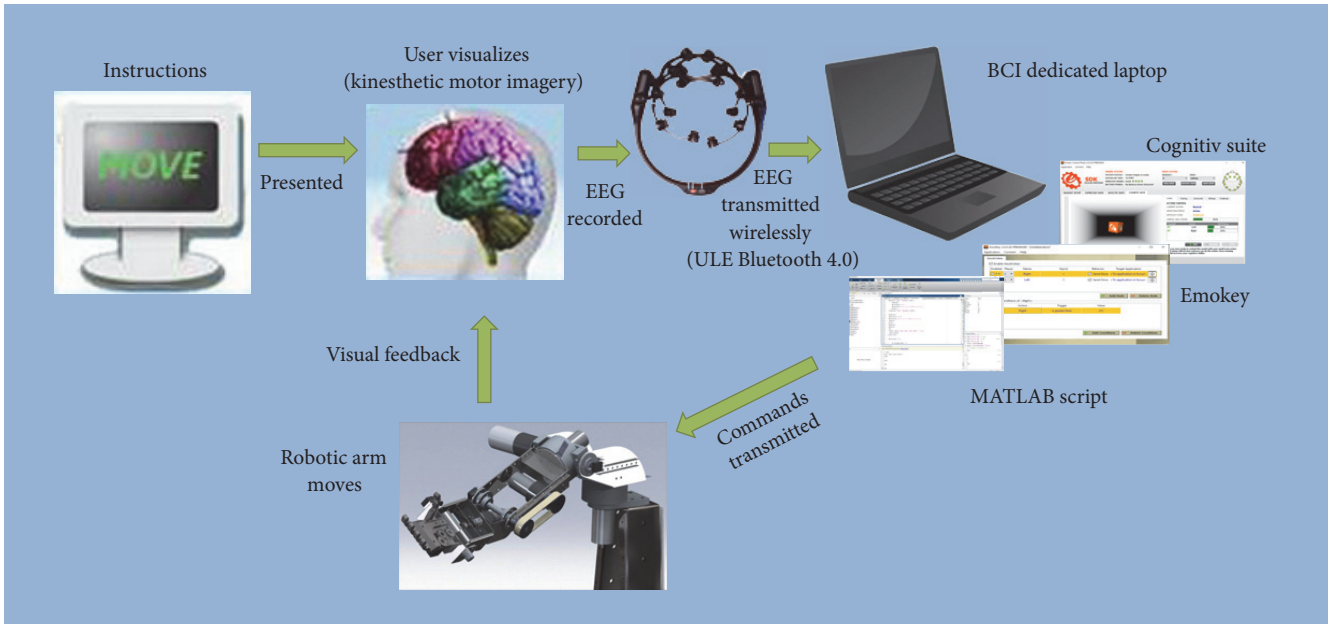


Figure 1: Brain-robot interface loop: using a wireless commercial EEG device for unobtrusiveness and simplicity of the system [12].

each arm, six of them being DC electric motors: two on the robotic shoulder joint, responsible for horizontal (1) and for vertical movement (2); one on the elbow joint, for horizontal movement (3); one on the wrist joint, for horizontal movement (4), one for rotation between the shoulder and elbow (5); and one for rotation between the elbow and wrist (6); finally, two servo motors add the ability of curling movement of the thumb and palm fingers (7) and (8). These 8 DoFs allow for responsive, fluid movement of the robotic arm and provide it with the ability to grab and manipulate small and light objects. Each arm functions with a single attached microcontroller, responsible for motors, connected to the dedicated computer system through a serial port at a Baud Rate of 9600. The microcontroller translates the commands the computer system transmits into positional coordinates for each motor. System responsiveness has been measured to be approximately 0.2 seconds.

2.2. Experimental Setup

2.2.1. Ethical Approval and Recruitment. The institutional bioethics committee approved the experimental protocol [12, 31] and all subjects filled and signed an informed consent form prior to their participation. Criteria for participation included clinical diagnosis and radiological documentation of SCI (evaluated by ASIA Impairment Scale (AIS) [33] and/or neurological examination reporting the condition) or healthy participants [31]. Exclusion criteria were other neurological injury or disease (traumatic brain injury, central nervous system tumors, epilepsy, etc.), recent participation in an interventional study, other grave medical condition that could affect participation or the safety, hearing and visual impairments, illegal drug use, and chronic alcoholism.

Table 1: Descriptive statistics of age (mean and standard deviation) for both groups (SCI group and healthy controls).

Group	Age Mean (std)
SCI	46.0 (17.64)
Healthy	46.2 (18.27)

Ten subjects with SCI (8 male, 2 female) and ten healthy controls (with an effort for being gender and age matched to the SCI group) were recruited in total. The SCI group had a mean age of 46.0 years (range 28–74, standard deviation (std) 17.64), while the healthy group had a mean age of 46.20 years (range 27–74, std 18.27), as shown in Table 1. The subjects had no prior experience in BCI or robotics.

2.2.2. Subject Assessment. For both groups demographics and medical history were collected including smoking habits, age, height, weight, Body Mass Index (BMI), and education level. BMI was calculated and the following classification was used: underweight < 18.5, normal weight $= 18.5$–24.9, overweight $= 25$–29.9, and obesity ≥ 30. For the SCI group, history data was also collected including age at injury, weight at injury, and cause of injury. Neurological examination was performed by a specialist physician using the International Standards for Neurological Classification of Spinal Cord Injury: severity of injury (classification in AIS), Neurological Level of Injury (NLI), motor scores for upper extremities (UEMS), lower extremities (LEMS), and total (Moto-Total), and sensory scores for light touch (LT), pin prick (PP), and total (Sensory-Total) were recorded [33]. Spasticity, if present, was also recorded using Ashworth [34] and Penn Spasm Frequency

Table 2: Subjects by group (SCI and healthy controls): basic demographic data for both groups and cause of injury, ASIA Impairment Scale (AIS), and Neurological Level of Injury.

SCI group	Age	Gender	Cause	AIS	NLI	Healthy group	Age	Gender
CSI-02-001	28	f	MVA	ASIA C	C4	CSI-04-001	27	f
CSI-02-002	52	m	MVA	ASIA D	C4	CSI-04-007	51	m
CSI-02-003	42	m	MVA	ASIA D	C8	CSI-04-003	43	m
CSI-02-004	70	m	Fall	ASIA D	C5	CSI-04-006	71	m
CSI-02-005	60	m	Fall	ASIA E	C6	CSI-04-009	63	m
CSI-02-006	28	m	MVA	ASIA D	C5	CSI-04-004	28	m
CSI-02-007	30	m	MVA	ASIA E	C5	CSI-04-005	31	m
CSI-03-001	47	m	Fall	ASIA A	T7	CSI-04-008	47	m
CSI-03-002	29	f	MVA	ASIA B	T4	CSI-04-002	27	f
CSI-03-003	74	m	Other	ASIA B	T4	CSI-04-010	74	m
Mean	46.00	-	-	-	-	*Mean*	46.20	-
Std	17.64	-	-	-	-	*Std*	18.27	-

Table 3: Neurological evaluation scores of the SCI group.

SCI group	Motor-Total	UEMS	LEMS	Sensory-Total	LT	PP
CSI-02-001	48	24	24	224	112	112
CSI-02-002	84	39	45	183	94	89
CSI-02-003	98	48	50	224	112	112
CSI-02-004	98	48	50	223	111	112
CSI-02-005	100	50	50	224	112	112
CSI-02-006	100	50	50	222	111	111
CSI-02-007	100	50	50	224	112	112
CSI-03-001	56	50	6	121	58	63
CSI-03-002	50	50	0	156	78	78
CSI-03-003	54	50	4	156	78	78
Median	91.00	50.00	47.50	222.50	111.0	111.50
Interquartile range	47.00	4.30	44.50	68.00	34.00	34.00

Scales [35]. Both groups were assessed for everyday life independence using the Spinal Cord Independence Measure (SCIM-III) [36], translated in Greek (g-SCIM-III) [37]. Total SCIM score and subscores for Self-Care (g-SCIM-III-SC), Respiration and Sphincter Control (g-SCIM-III-RS), and Mobility (g-SCIM-III-M) were recorded.

In the SCI group, 9 out of 10 patients had incomplete injury. Four patients were classified as AIS D (40%), 2 were AIS B (20%), 2 were AIS E (20%), one had complete injury and was classified as AIS A (10%), and one patient was classified as AIS C (10%). Regarding level, 70% were cervical injuries whereas the remaining 30% were thoracic injuries. Moreover, the predominant cause of injury, in 60% (6/10) of patients injury, was motor-vehicle accidents (MVA), while in 30% of participants the injury was induced by fall from heights (Fall) and one patient reported other causes. Table 2 depicts age and gender distributions for both groups and also cause of injury, AIS, and NLI by subject.

Based on neurological assessment, the 60% of patients that were classified as either AIS D or E showed approximately intact general motor function and excellent UEMS and LEMS (3/10 of patients scored 50 in both categories whereas 2/10 scored 48 in UEMS and 50 in LEMS). The remaining 40% patients were classified as AIS A, B, or C and showed motor deficits, as presented in Table 3. The SCI group was therefore further grouped into positive outcome (60%) and negative outcome (40%) for further analysis as described in the statistical analysis section below. With regard to sensory skills, the patients with good outcome scored as high as the healthy controls in LT and PP sensory examination (Table 3).

Subject assessment also included subjective reporting of imagery capacity, using Vividness of Visual Imagery Questionnaire (VVIQ-II) [38] with eyes open that was assessed for total score and for each of the four scenarios (VVIQ1-VVIQ4). Psychometric evaluation also entailed answering Beck Depression Inventory (BDI) [39] and Rosenberg Self-Esteem Questionnaire (RSEQ) [40], both translated in Greek [41, 42]. After the participation in the experiment, both subject groups reported on HRI experience using the Godspeed robotics questionnaire [27], also translated in Greek [43]. Godspeed total score (GDSPD-Total) and subscores for Anthropomorphism (GDSPD-Anthr), Animacy (GDSPD-Anim), Likeability (GDSPD-Like), Perceived Intelligence (GDSPD-Int), and Perceived Safety (GDSPD-Safe) were recorded.

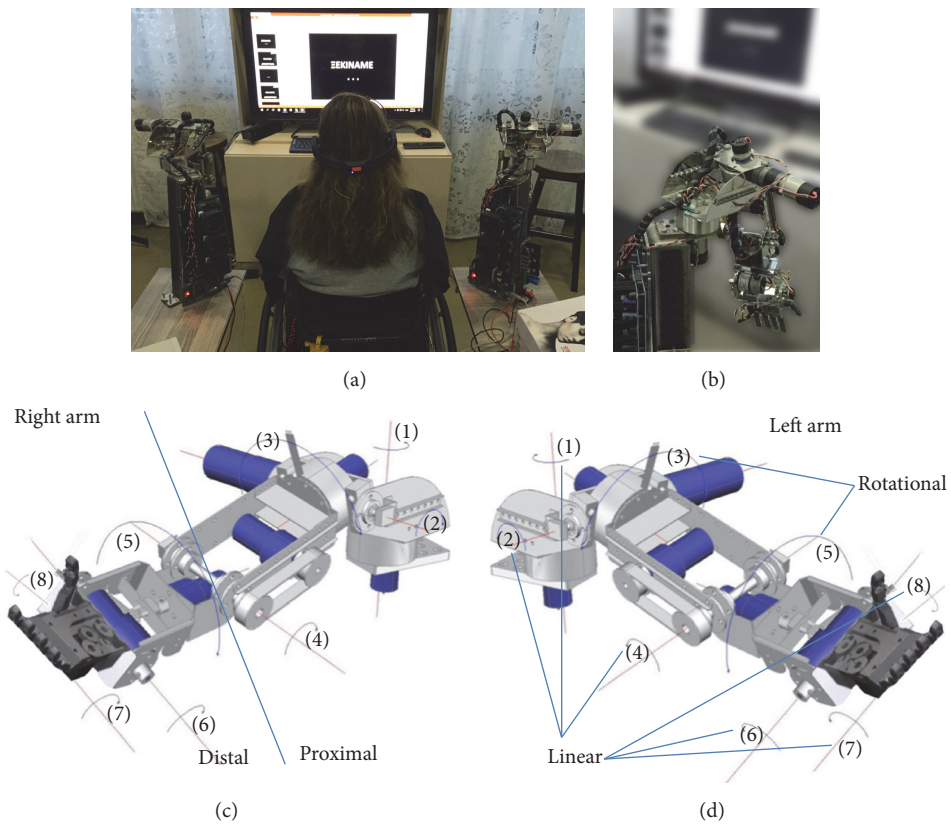


Figure 2: The experimental setup in Thess-AHAL. (a) SCI subject seated across the TV/computer monitor and between two robotic arms, wearing a wireless commercial EEG device; (b) close-up to a "Mercury 2.0" house-built robotic arm; (c) the right robotic arm, showing 8 possible DoFs, grouped into proximal and distal movements; (d) the left robotic arm, showing DoFs that result in rotational or linear movement. Each DoF allows movement towards two possible directions [12].

2.2.3. Experimental Procedure. As the experimental procedures have been also described in detail [12, 29], we will provide an overview hereby. The subjects were asked to perform Visual Motor Imagery (VMI), meaning they had to mentally rehearse 32 different movements of the upper extremities, while watching a set of arms performing randomly those movements on a computer screen. Each movement was shown 9 times in total. Randomly, walking and an oddball video were also shown. This VMI experience was aimed at preparing participants for the BCI classes they would have to later perform. This VMI procedure was also performed under high-resolution EEG recording, which will not be further elaborated hereby. The subjects had their arms, torso, and legs covered with a black curtain during VMI experience as well as during subsequent BCI training and BCI control to facilitate registration of the projected arms (in the case of VMI) or the robotic arms (in the case of BCI) into the body schema [28]. Also, in the case of VMI the projected arms were gender-matched to the subject.

Following VMI experience, the subjects sat across a computer monitor, between two Mercury 2.0 robotic arms, located in the Thessaloniki Active and Healthy Ageing Living Lab technology showcase room (Thess-AHALL, member of ENoLL) [44–46]. Subjects wore an Emotiv EPOC headset and they began training of the BCI modality using kinesthetic motor imagery (KMI) of movements of the left and the right arms. Action power threshold was set at 50% and all subjects were above threshold during training [29]. Three imagery classes were trained (resting state, left, and right). Five training cycles were allowed for left and right, using continuous mental rehearsal of all movements of left and right arm, respectively (as in the videos that subjects watched) [47]. Training skills (Train-L) (Train-R) classes were recorded. Following BCI training, the subjects were given written commands in the monitor to move a specific DoF towards a specific direction. 32 different commands were given in total and the subjects were asked to use the trained KMI skill in order to control the robotic arms to perform those actions. Each command was projected for 30 sec with a 5 sec rest between them (Figure 2). During the 30 sec period the subject attempted to perform the movement as many times as possible by giving the desired direction of imagery class (left or right). The BCI detected as correct (desired) any activation of the class corresponding to the arm currently targeted and gave the output to move the arm as per the instruction (positive feedback). Incorrect (undesired) activation that was detected resulted in an opposite motion of the target robotic arm (negative feedback). Desired and undesired commands that were detected by the BCI program were recorded through the MATLAB script.

Table 4: House-developed performance rating scale for brain-robot interface control, tailored to the needs of the use case of the current experiment.

Performance rating of brain-robot interface control based on ratio of desired to undesired detected & classified mental commands		
Score	Rating	Description of rating
5	Excellent	Arms were moving towards desired direction for most of the time
4	Very Good	5+ commands were detected, minimal undesired commands
3	Good	4+ commands were detected, only 2-3 undesired commands
2	Average	3+ commands were detected, but also up to +3 undesired commands
1	Minimal	1-2 commands were detected *or* 3-4 including undesired commands *or* many commands were detected but most were undesired
0	No control	No command *or* only undesired commands were detected

Performance was rated in each different movement with a score from 0 to 5, based on the ratio of desired mental commands to undesired mental commands that were detected and classified (Table 4) [12]. A total score for BCI performance (BCI-total) was then calculated for each subject, adding the scores for each movement (max 160) and the performance of each subject was also converted to a percentage (BCI%). Subscores were calculated for the 16 movements of the left arm (BCI-L), the 16 movements of the right arm (BCI-R), the 16 movements of wrists, fingers, and thumbs (BCI-Distal), and the 16 movements of shoulder and elbow (BCI-Proximal), with max scores of 80. Subscores were also calculated for 24 linear movements and 8 rotational movements but, to allow comparison, an average per movement score was calculated for both categories (BCI-Li/24 and BCI-Ro/8).

2.3. Statistical Analysis

2.3.1. Statistical Tests. Statistical analysis was performed in IBM SPSS Statistics (version 23) and we set a significance level of 0.05 for all statistical tests. The variables were explored for normality assumption using two grouping schemes: (1) SCI and healthy groups as grouping factor for all subjects and (2) positive and negative outcome as grouping factor for SCI subjects.

Normality was explored using visual inspection of histograms, *normal Q-Q plots* and *boxplots*, *skewness*, and *kurtosis* [48–50] as well as using the normality tests (*Shapiro-Wilk test* and *Kolmogorov-Smirnov Test*) [51, 52]. The variable age, in particular, was normally distributed for both groups (*skewness:* 0.407 (SE = 0.687), *kurtosis:* −1.418 (SE = 1.334) *for healthy group; skewness:* 0.651 (SE = 0.687), *kurtosis:* −0.752 (SE = 1.334) *for patient group*).

Normality assumption was not met (1) for motor and sensory scores of neurological examination as well as SCIM scores, (2) for VVIQ1 and VVIQ3 scenarios and BDI in SCI and healthy groups, (3) for GDSPD-Safe in either grouping scheme, and (4) for training scores in outcome grouping. All other distributions met normality assumption. Depending on normality assumption group differences were explored using either *Independent Samples t-test* or *Mann–Whitney (U) test*. Possible associations between quantitative variables were explored via *Pearson correlation coefficient* or *Spearman's coefficient* depending on normality. Other specific statistical tests or other specific groupings were used as described below. Please also refer to Supplementary Materials for a more detailed report of statistical analysis (available here).

2.3.2. Demographics, Somatometric Data, and Clinical Evaluation. Initially, we planned comparisons of demographic and somatometric data between SCI and healthy groups. Comparisons regarding the education level (basic studies, pregraduate level, graduate level, postgraduate level, and Ph.D. holders) were performed by *Mann–Whitney (U) test*, since our data are ordinal (Likert-type) [53]. Medical history and neurological data were analyzed via descriptive statistical methods. Moreover, we explored whether the smoking status is independent of the group or not using *chi-square test*.

2.3.3. Assessment Questionnaires and User Perception. Imagery capacity (VVIQ) and psychometric questionnaires (BDI and RSEQ) were analyzed between SCI and healthy groups, as well as between SCI subgroups of positive and negative outcomes. The scores of Godspeed and its subcategories (GDSPD-Total, GDSPD-Anthr, GDSPD-Anim, GDSPD-Like, GDSPD-Int, and GDSPD-Safe) were analyzed between patient and healthy groups, as well as between patient outcome groups. Finally, possible correlations were explored between Godspeed scores and BCI performance as well as Godspeed scores and VVIQ scores.

2.3.4. BCI Performance. BCI-total, BCI-L, BCI-R, and training skills (Train-L and Train-R) were analyzed between SCI and healthy groups as well as in SCI outcome subgroups. BCI-Distal, BCI-Proximal, BCI-Li, and BCI-Ro were analyzed using descriptive statistics. We further explored differences in BCI performance and training scores between different groups of neurological levels of injury (cervical, thoracic) after testing for normality. Additionally, linear regression analysis was used to model the possible relationship between the independent variable BCI scores and the explanatory variable NLI using linear regression analysis. Possible correlations were further explored across groups (SCI/healthy) between BCI performance and (1) age, (2) imagery capacity (VVIQ), and (3) psychometric questionnaires (BDI, RSEQ) for both groups.

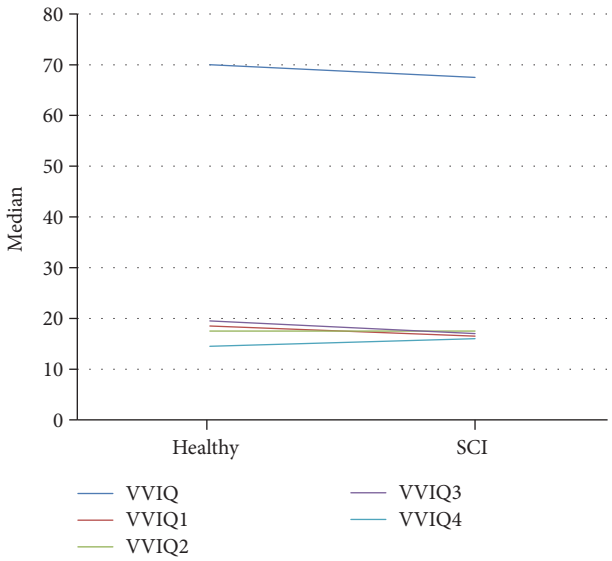


Figure 3: Scores of SCI and healthy groups at VVIQ test and its scenarios show no significant difference between groups.

3. Results

3.1. Demographics, Somatometric Data, and Clinical Evaluation. Planned comparisons of age between groups did not reveal any significant difference (Table 1). Regarding education level, no statistically significant difference between groups was shown. Group differences were not revealed in somatometric data either. Most of the injuries (4/10) happened when the participants were between 26.0 and 33.0 years of age. The second more vulnerable age range to injuries (MVAs) seems to be the age range of 50.0–57.0 years. Exploring whether the smoking status is independent of the group or not, we observed a significant association ($\chi^2 = 5.051$, df = 1, $\mathbf{p} = \mathbf{0.025}$), since 70% of the healthy group and only 20% of the SCI group were smokers. Subjects with positive outcome scored as high as the healthy subjects in g-SCIM-III test and its subcategories. In the remaining group, deviation in the performance at the g-SCIM-III test was found. Table 5 shows aggregated the most important statistical tests we performed and their results on demographics, clinical evaluation, and assessment questionnaires.

3.2. Assessment Questionnaires. VVIQ did not show any significant difference between SCI and healthy groups in total score or any of the scenarios. SCI group scored lower than healthy (Figure 3) but differences did not reach statistical significance. Regarding SCI outcome subgroups, no considerable group differences were revealed in VVIQ scores. Also, even though the SCI group showed increased depressive scores and decreased self-esteem compared to healthy group, group differences were not statistically significant (Figure 4).

3.3. BCI Performance. Planned between-group (SCI/healthy) comparisons of BCI performance revealed a statistically significant difference only in the control of right robotic arm (BCI-R: $t = 2.592$, df = 18, $\mathbf{p} = \mathbf{0.018}$). Although BCI scores of SCI group were generally lower than those of control, group differences did not reach statistical significance. Also, regarding performance in different types of robotic arm movements (Table 6), based on the planned analysis, no significant outcomes were shown. Comparisons of BCI scores between SCI outcome subgroups (positive, negative) did not reveal any statistically significant difference (Figure 5). However, subgroup with positive outcome seems to score higher in total BCI control (Figure 6).

The ability of SCI subjects to control robotic arms did not differentiate depending on the injury location (cervical, thoracic). However, subjects with cervical injuries scored higher than those with thoracic injuries in BCI control. BCI-Total was marginally negatively correlated to NLI in the SCI group ($\mathbf{p} = \mathbf{0.064}$) (Figure 7). Negative correlations, but not statistically significant, were found between the BCI performance and the age (Table 6).

The training skill was not statistically different between SCI and healthy groups neither for the right hand, nor for the left hand, even if SCI subjects showed slightly lower training scores in the right hand and higher in the left hand than healthy participants. Comparing training scores between SCI subjects grouped by outcome (Figure 8), a marginally considerable difference was found only in training scores of left hand ($U = 3.5$, $\mathbf{p} = \mathbf{0.068}$). SCI subjects with negative outcomes trained more efficiently in left hand compared to the subjects with positive outcome (Table 6). SCI subjects did not considerably differentiate their training scores depending on the level of injury. However, it was noted that Train-L was greater in thoracic SCI subgroup than those with cervical injuries whereas the opposite was the case for Train-R.

Total BCI scores were significantly negatively associated with VVIQ total scores ($r = -0.727$, $\mathbf{p} = \mathbf{0.017}$) and VVIQ3 ($r_s = -0.948$, $\mathbf{p} < \mathbf{0.001}$) as well as the BCI-L scores and VVIQ3 ($r_s = -0.665$, $\mathbf{p} = \mathbf{0.036}$) only in SCI group. However, other subcategories of VVIQ such as VVIQ1, VVIQ2, and VVIQ4 were not correlated to all BCI scores in SCI group (*BCI-Total-VVIQ1:* $r_s = -0.489$, $p = 0.151$; *BCI-L-VVIQ1:* $r_s = -0.509$, $p = 0.133$; *BCI-R-VVIQ1:* $r_s = -0.194$, $p = 0.590$; *BCI-Total-VVIQ2:* $r = -0.077$, $p = 0.832$; *BCI-L-VVIQ2:* $r = -0.073$, $p = 0.840$; *BCI-R-VVIQ2:* $r = 0.005$. $p = 0.989$; *BCI-Total-VVIQ4:* $r = -0.312$, $p = 0.380$; *BCI-L-VVIQ4:* $r = 0.026$, $p = 0.943$; *BCI-R-VVIQ4:* $r = -0.332$, $p = 0.348$). With regard to BCI performance depending on the depressive symptomatology as assessed by BDI, significant negative correlation was found only between total BCI scores and BDI scores in healthy participants ($r_s = -0.719$, $\mathbf{p} = \mathbf{0.019}$). Correlations explored between all BCI scores and scores at Rosenberg Self-Esteem Scale did not reach statistical significance for both groups (*Healthy-BCI-Total-Rosenberg:* $r = 0.150$, $p = 0.679$; *BCI-L-Rosenberg:* $r = 0.054$. $p = 0.882$; *BCI-R-Rosenberg:* $r = 0.115$, $p = 0.751$; *SCI: BCI-Total-Rosenberg:* $r = -0.067$, $p = 0.854$; *BCI-L-Rosenberg:* $r = -0.397$, $p = 0.255$; *BCI-R-Rosenberg:* $r = 0.369$, $p = 0.294$).

3.4. User Perception. Based on planned analysis of HRI characteristics of the robotic arms, SCI and healthy groups did not present significant differences in their answers in any subcategory (*GDSPD-Anthr:* $t = 1.504$, df = 18, $p = 0.150$;

Table 5: Statistical tests on demographics, clinical evaluation, and assessment questionnaires.

Demographics	SCI versus healthy		Clinical evaluation	SCI	Healthy
Education	$U = 33.50$	$p = 0.179$	g-SCIM-III-Total	Median = 94.0, IQR = 51.3	Median = 100.0, IQR = 0.0
Height	$t = 1.634$, df = 18	$p = 0.120$	g-SCIM-III-SC	Median = 18.0, IQR = 7.3	Median = 20.0, IQR = 0.0
Weight	$t = 1.177$, df = 18	$p = 0.254$	g-SCIM-III-BS	Median = 40.0, IQR = 19.3	Median = 40.0, IQR = 0.0
BMI	$t = 0.646$, df = 18	$p = 0.526$	g-SCIM-III-M	Median = 37.5, IQR = 24.3	Median = 40.0, IQR = 0.0
Smoking	$\chi^2 = 5.051$, df = 1	**p = 0.025**			

	Assessment questionnaires					
	SCI	Healthy	SCI versus healthy		Positive versus negative outcome	
VVIQ	65.70 (7.04)	67.0 (10.92)	$t = 0.316$, df = 18	$p = 0.755$	$t = -1.094$, df = 8	$p = 0.306$
VVIQ1	Median = 16.5, IQR = 4.0	Median = 18.5, IQR = 4.5	$U = 42.0$	$p = 0.533$	$t = 0.245$, df = 8	$p = 0.813$
VVIQ2	17.20 (2.25)	16.50 (2.99)	$t = -0.591$, df = 18	$p = 0.562$	$t = -1.884$, df = 8	$p = 0.096$
VVIQ3	Median = 17.0, IQR = 4.5	Median = 19.5, IQR = 3.3	$U = 29.50$	$p = 0.113$	$t = 0.086$, df = 8	$p = 0.934$
VVIQ4	15.10 (3.98)	14.70 (3.59)	$t = -0.236$, df = 18	$p = 0.816$	$t = -1.233$, df = 8	$p = 0.253$
BDI	Median = 14.5, IQR = 12.5	Median = 3.0, IQR = 5.5	$U = 26.0$	$p = 0.069$		
RSEQ	21.40 (3.86)	24.70 (4.14)	$t = 1.843$, df = 18	$p = 0.082$		

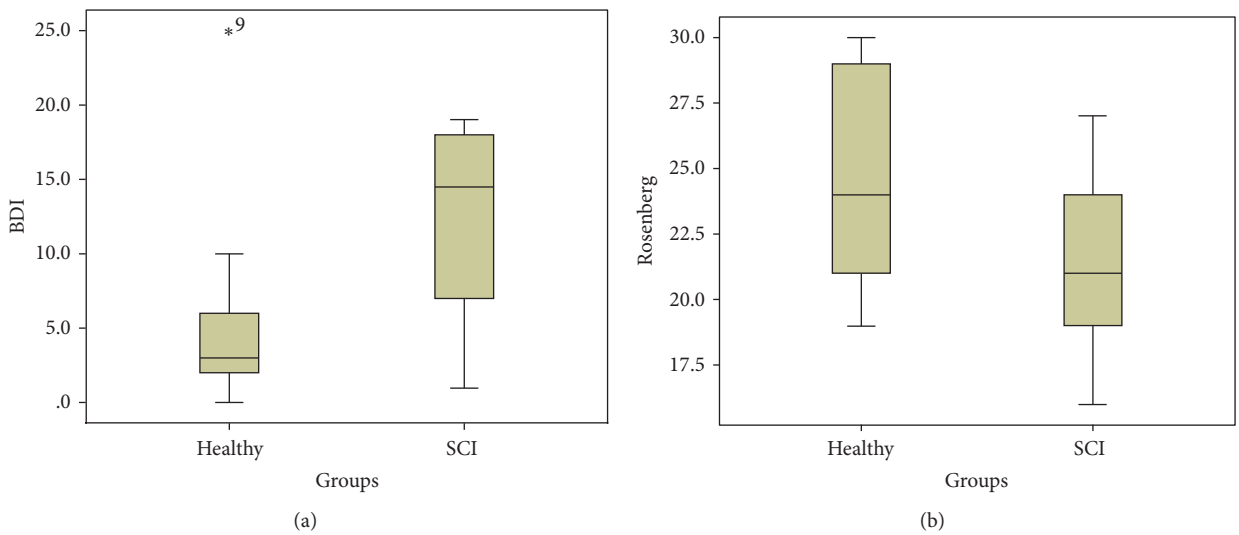


Figure 4: SCI scores at Beck Depression Inventory are increased while at Rosenberg Self-Esteem Scale they are decreased, compared to healthy controls, but findings are not statistically significant. *Extreme value.

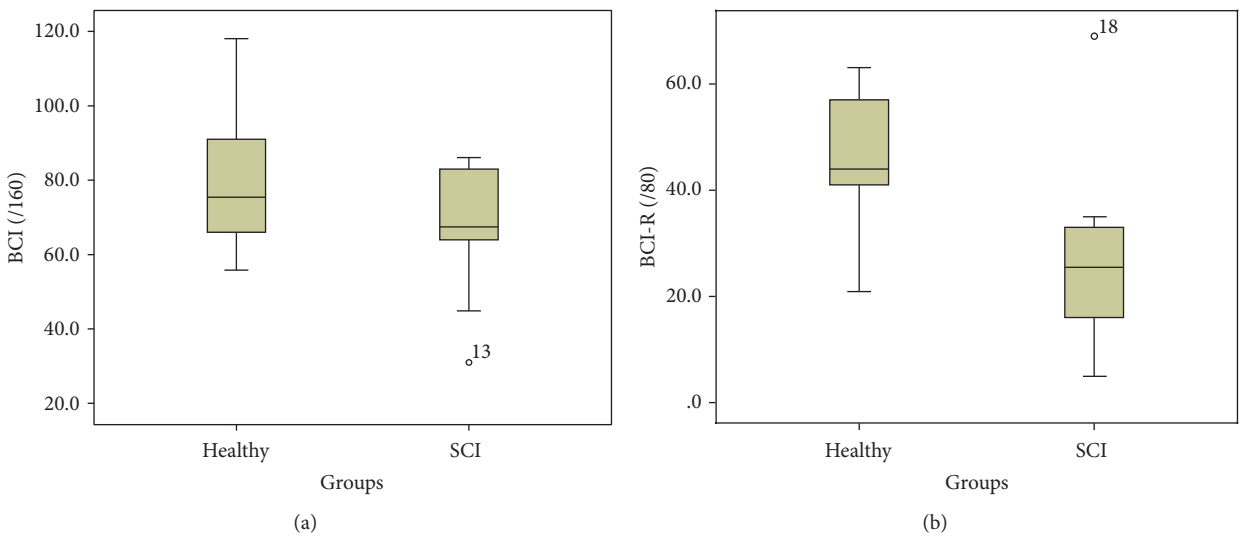


Figure 5: BCI performance (BCI-Total) in the control of robotic arms and in the control of the right robotic arm (BCI-R) was lower for SCI group, but not statistically significant. ∘: outlier.

GDSPD-Anim: $t = 0.611$, df = 18, $p = 0.549$; *GDSPD-Like:* $t = -0.217$, df = 18, $p = 0.831$; *GDSPD-Int:* $t = -0.047$, df = 18, $p = 0.963$; *GDSPD-Safe:* $U = 42.0$, $p = 0.536$) of Godspeed questionnaire. Additionally, there was no difference between groups in GDSPD-Total ($t = 0.427$, df = 18, $p = 0.675$).

The scores are illustrated in Figure 9 (GDSPD (/120) (mean (SD)), healthy: 80.80 (14.78); SCI: 78.00 (14.57); GDSPD-Anthr, healthy: 13.60 (4.25); SCI: 10.80 (4.08); GDSPD-Anim, healthy: 18.90 (5.11); SCI: 17.70 (3.53); GDSPD-Like, healthy: 18.50 (3.87); SCI: 18.90 (4.36); GDSPD-Int, healthy: 17.70 (4.60); SCI: 17.80 (4.92); GDSPD-Safe, healthy (median = 12.0, IQR = 5.3); SCI (median = 13.50, IQR = 4.5)).

Godspeed's scores were further explored after grouping by outcome (positive, negative). Marginally significant difference was found only in Anthropomorphism ($t = 2.251$, df = 8, $p = \mathbf{0.054}$) *(GDSPD (/120):* $t = 1.918$, df = 8, 0.091; *GDSPS-Anim:* $t = 1.382$, df = 8, $p = 0.204$; *GDSPD-Like:* $t = 1.289$, df = 8, $p = 0.233$; *GDSPD-Int:* $t = 1.343$, df = 8, $p = 0.216$; *GDSPD-Safe:* $U = 10.50$, $p = 0.741$*)*. More precisely, SCI subjects with negative outcome scored higher than those with positive outcome in this Godspeed's subcategory *(negative outcome: 13.75 (3.30); positive outcome: 8.83 (3.43))*.

Significant correlation between Godspeed's and BCI scores was not revealed for any group. GDSPD-Anthr and GDSPD-Int were positively correlated to VVIQ4 and VVIQ1

Table 6: Statistical tests on BCI performance and training skills.

			BCI performance			
	SCI	Healthy	SCI versus healthy		Cervical	Thoracic
BCI-Total	66.70 (17.49)	79.50 (18.51)	$t = 1.589$, df = 18	$p = 0.129$	71.57 (14.75)	55.33 (21.13)
BCI (%)	41.69% (10.93%)	49.65% (11.47%)	$t = 1.590$; df = 18	$p = 0.129$		
BCI-R	27.10 (17.97)	45.80 (14.05)	$t = 2.592$, df = 18	**p = 0.018**	29.86 (20.07)	20.67 (12.50)
BCI-L	39.60 (19.64)	33.70 (21.32)	$t = -0.644$, df = 18	$p = 0.528$	41.71 (19.20)	34.67 (24.03)
BCI-Distal	33.40 (9.98)	40.50 (8.45)	$t = 1.717$, df = 18	$p = 0.103$		
BCI-Proximal	33.30 (8.82)	39.00 (11.95)	$t = 1.213$, df = 18	$p = 0.241$		
BCI-Linear/24	2.11 (0.52)	2.48 (0.60)	$t = 1.286$, df = 18	$p = 0.215$		
BCI-Rot/8	2.25, 1.06	2.44, 0.72	$U = 33.00$	$p = 0.196$		
	Positive	Negative	Positive versus negative outcome		Cervical versus thoracic	
BCI-Total	69.50 (15.00)	62.50 (22.43)	$t = -0.598$, df = 8	$p = 0.567$	$t = 1.420$, df = 8	$p = 0.194$
BCI (%)	43.44% (9.36%)	39.06% (14.02%)	$t = -0.597$, df = 8	$p = 0.567$		
BCI-R			$t = -0.390$, df = 8	$p = 0.707$	$t = 0.721$, df = 8	$p = 0.491$
BCI-L			$t = -0.168$, df = 8	$p = 0.871$	$t = 0.498$, df = 8	$p = 0.632$
			Training skill			
	SCI	Healthy	SCI versus healthy		Cervical	Thoracic
Train-R	46.30% (18.00%)	46.40% (10.32%)	$t = 0.015$, df = 18	$p = 0.988$	49.14% (18.05%)	39.67% (19.66%)
Train-L	56.20% (19.61%)	53.80% (16.67%)	$t = -0.295$, df = 18	$p = 0.771$	53.71% (22.54%)	62.00% (11.53%)
	Positive	Negative	Positive versus negative outcome		Cervical versus thoracic	
Train-R	Median = 48.00%, IQR = 40.25%	Median = 50.50%, IQR = 26.50%	$U = 11.50$	$p = 0.915$	$t = 0.744$, df = 8	$p = 0.478$
Train-L	Median = 51.00%, IQR =19.00%	Median = 68.00%, IQR = 40.00%	$U = 3.5$	**p = 0.068**	$t = -0.590$, df = 8	$p = 0.572$
			Correlations			
	NLI		Age-SCI		Age-healthy	
BCI-Total	$\beta = -2.656$, CI [−5.512, 0.2], **p = 0.064**, adjusted R square = 0.286		$r = -0.507$	$p = 0.135$	$r = -0.361$	$p = 0.305$
BCI-R	-	-	$r = -0.239$	$p = 0.506$	$r = -0.294$	$p = 0.410$
BCI-L	-	-	$r = -0.233$	$p = 0.517$	$r = -0.120$	$p = 0.741$

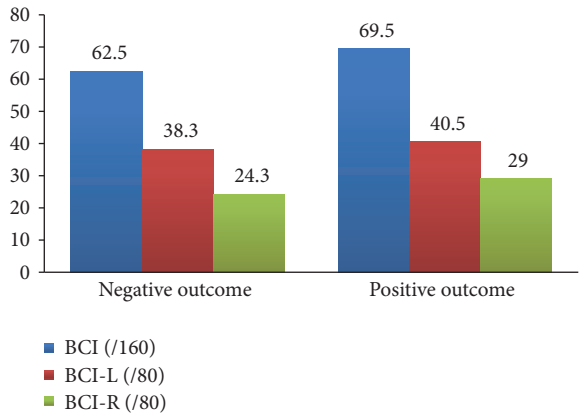


Figure 6: BCI scores of SCI subjects grouped by neurological outcome. Subjects with positive outcome performed nonsignificantly better in BCI control than those with negative.

scenario's scores, respectively, only in healthy (*GDSPD-Int, VVIQ4:* $r = 0.654$, $p = \mathbf{0.040}$*; GDSPD-Anthr, VVIQ1:* $r = 0.629$, $p = \mathbf{0.052}$). In SCI group, GDSPD-Safe was positively associated with scores in VVIQ total ($r = 0.696$, $p = \mathbf{0.025}$) and VVIQ1 ($r = 0.780$, $p = \mathbf{0.008}$).

4. Discussion

4.1. User Perception and Performance Assessment. Our results show that healthy controls performed better (49.65%) than the SCI group (41.65%) in BCI control and so did subjects with positive neurological outcome (43.44%) when compared to subjects with negative neurological outcome (39.6%). Both those observations did not reach statistical significance but their interpretation is not as straightforward. Since no patients with complete tetraplegia were included, these findings probably cannot be attributed to some type of "thought extinction process" [54]. Indeed, self-reported imagery capacity (as measured by VVIQ) was found to be nonsignificantly lower in SCI group than in healthy subjects but, on the other hand, subjects with cervical injuries fared better than those subjects with thoracic injuries. Thus the group with injury that could affect upper extremity neural circuits outperformed those subjects that did not have direct injury to those circuits. Moreover this finding was further reinforced by a marginally significant (negative) correlation of BCI performance to neurological level of injury, meaning that the higher the level of injury, the better the subjects performed.

Possible explanations for performance differences could be explored along the lines of depression [55] and motivation [56, 57]. Indeed the SCI group showed increased depressive scores (measured by BDI) and lower self-esteem scores (measured by RSEQ), despite both findings not being statistically significant. Our research protocol did not include any official questionnaire to assess SCI and healthy subject with regard to motivation. Only by anecdotal evidence, during informal debriefing discussions that we held with participants, subjects with thoracic injuries tended to consider somewhat irrelevant this experimental setup to their condition. While reciprocal sensorimotor pathways and sensorimotor cortex representations are affected regardless of the level of SCI, those subjects felt that their clinical condition and independence demands were not addressed by robotic arms technology.

The HRI characteristics of the robot were measured (after the participation) by Godspeed robotics questionnaire. It seems that user perception did not differ significantly between SCI and healthy subjects. While in Figure 9 we show absolute values for the total questionnaire score and its subscores, since no standard for direct comparison of those can be identified, we feel that intergroup comparisons are more useful. As such, we believe that this is a positive finding, interpreting it as users with SCI appreciating the robot more or less the same as their healthy counterparts. Among SCI subgroups, those with negative outcome gave higher scores to Anthropomorphism of the robot (mean 13.75 out of max 25) than those with positive outcome (mean 8.83 out of max 25), a finding that can be considered marginally significant (see Figure 9(b)). Indirectly, since user perception did not correlate with BCI performance and did not vary between SCI and healthy subjects in all other comparisons, perhaps this could be attributed to psychometric attributes of the participants and it should be further investigated in an SCI group with wider participation.

As we mentioned, no statistically significant negative correlations were found between BCI performance and the age of the subjects, a finding that also needs further study in order to be validated. As much of the BCI research field endeavors have been shaped around the issue of providing affordable, acceptable, and useful assistive systems for the benefit of the disabled, the particular characteristics of elderly patients should be the focus of investigation [58, 59].

Finally, BCI performance was significantly negatively associated ($p = 0.017$) with imagery capacity (as self-reported by VVIQ), a seemingly nonsensical finding. That could probably be attributed to one of the scenarios (VVIQ3 $p < 0.001$) skewing the total. We have insufficient data to interpret this, but until further evidence our hesitant (since this is an established tool [60]) suggestion would be to either to use a different research tool to measure imagery capacity or to omit the 3rd scenario of the VVIQ in the context of BCI controlled robotic arm experiments.

Some correlation was also revealed between reported imagery capacity and Godspeed questionnaire. SCI subjects that had higher scores in VVIQ rated Perceived Safety higher ($p = 0.025$). This finding was strongly influence by VVIQ 1st scenario ($p = 0.008$) since it was not present for the other scenarios. Said scenario asks of the subjects to think of some relative or friend whom they frequently see and consider their picture, a mental task that could induce a sense of familiarity or attachment and perhaps affect their answer on the Perceived Safety questions [61]. On the other hand, healthy subjects that had higher scores in VVIQ 1st and 4th scenarios rated Perceived Intelligence and Anthropomorphism higher too, findings that are not explained by the content of those scenarios. Perhaps their significance is marginal ($p = 0.052$ and $p = 0.040$, resp.) but whether these trends persist or not with a wider user base is something that should be investigated before drawing an accurate conclusion.

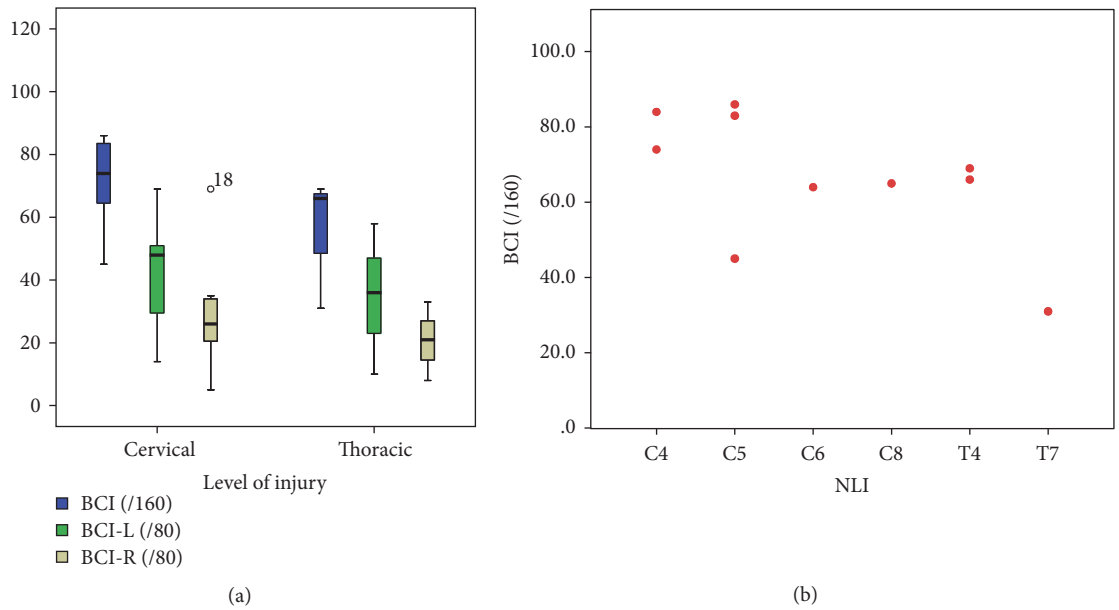


Figure 7: BCI performance of SCI group depending on the level of injury (a) and significantly (marginal inversely) correlated to NLI. ∘: outlier.

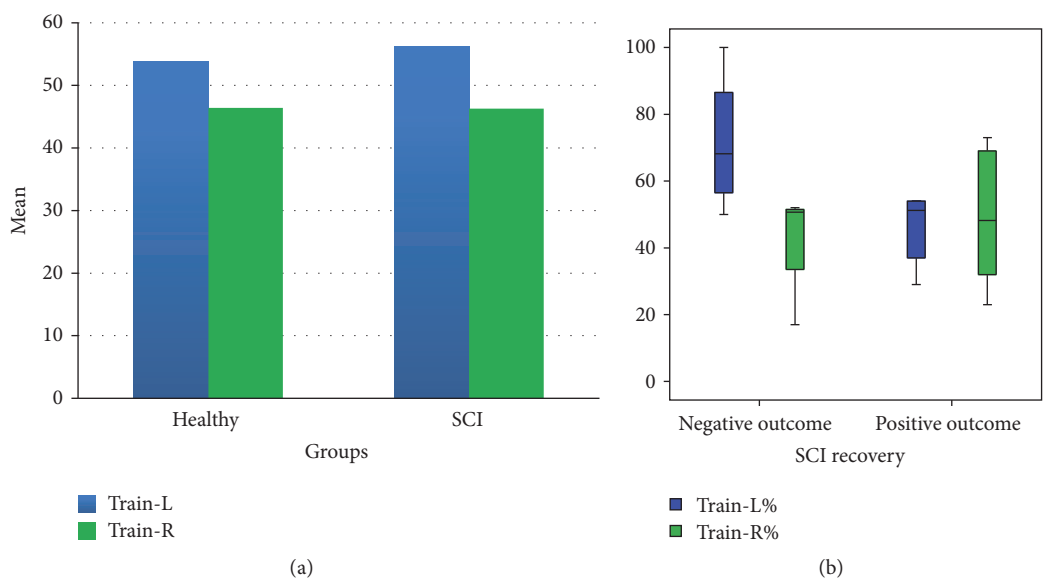


Figure 8: Train-L and Train-R, mean values—grouped by SCI or healthy group (a) and median values—grouped by outcome (b). SCI subjects with negative outcome performed marginally better than those with positive.

Our work adds onto an envisaged puzzle of BCI technology and uses case scenarios with an emphasis on affordability (off-the-shelf component system) and realism at a reasonable accuracy. Multi DoF robotics control by users with SCI has been proven to be feasible even by commercial-class wireless BCI. Nonetheless, repeated training in several BCI sessions is probably needed to improve performance by the end-users. Although users with complete lesions by definition would benefit more from ongoing research in robotics and BCIs, the use of brain controlled assistive technologies holds special importance for users with incomplete injuries too [62]. As a result of preserved reciprocal neural pathway communication, robotic-assisted training and brain controlled functional stimulation have been shown to improve arm motor function of those patients [63, 64] but expectations from assistive technologies and motivation to use them should be among key targets for further user experience investigations.

4.2. Further Technological Development. Our main development directions include (1) improving accuracy, speed, responsiveness, fluidity, and efficiency; (2) enhancing the integration of the system into the operator's perceived body

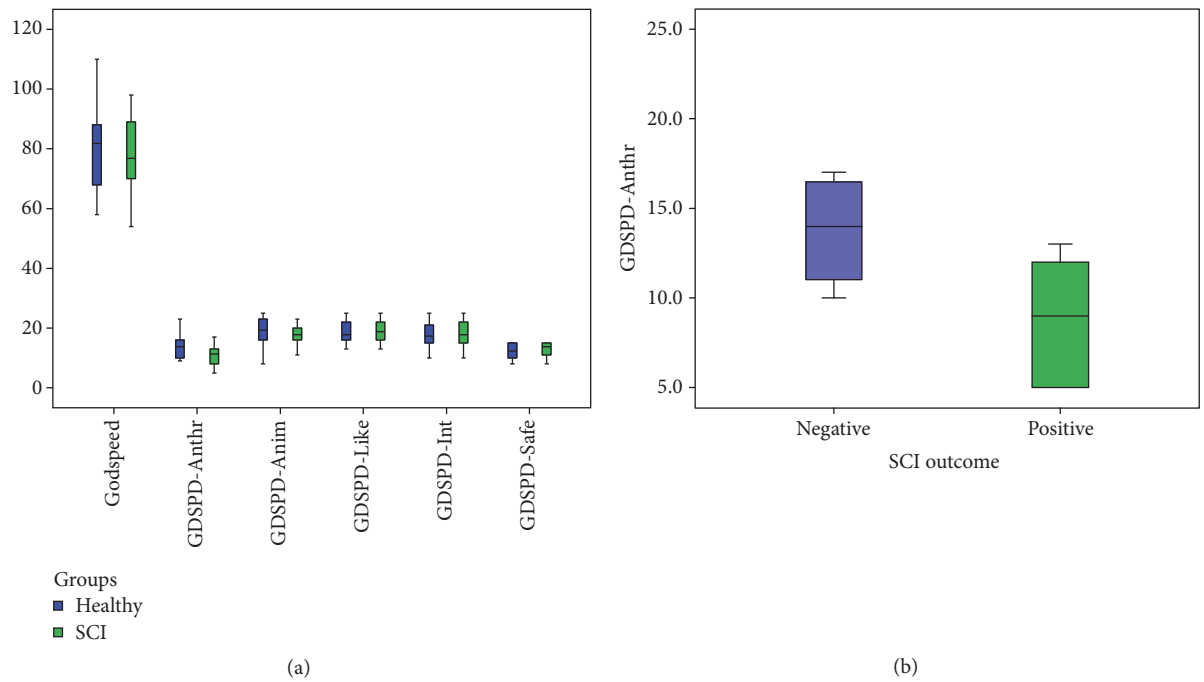


Figure 9: Answers provided to all Godspeed scores from SCI and healthy groups (a) showing no significant difference between groups. Anthropomorphism (b) was rated marginally higher by SCI subjects with negative outcome than those with positive.

mental image [17]. The pipeline that we overview below is tested offline, since several aspects, such as the solution of the inverse problem, entail high computational cost and are planned to be realized with dedicated processing units and use of artificial intelligence (AI) [12]. We are currently developing and testing a novel BCI decoding algorithm. Our goal is to increase discriminative ability to multiple MI classes projecting sensor data to cortical source domain [65, 66]. The COLIN27 anatomy is used [67], but tests are also made on individual head models. Our approach is based on studies indicating that the features extracted from the source domain enhance the discrimination [68–70], since the higher dimensional space of sources incorporates anatomical and physiological information. Spatial features are extracted in predefined Regions of Interest (ROIs) [12] and Common Spatial Pattern (CSP) filters [71] are calculated on each ROI. Tuning and feature selection using a linear classifier (LDA) are a crucial part in order to determine the most valuable ROIs and improve BCI performance. The use of time-frequency and connectivity features in conjunction will also be explored.

4.3. Limitations. As we have already mentioned [12], multiple MI class discrimination is a challenge that state-of-the-art noninvasive BCI technology has yet to meet with success. Depending on a proprietary BCI algorithm, such as in our case, further underlines this problem, as it was not tailored to our specific needs. While the commercial-class device meets the use case needs and will be retained in future implementations, the proprietary BCI algorithm will be substituted with the one currently under development, in order to explore and reach practically usable performance of the entire system. This setup will be used in future experiments, including tests on the EEG data that have been already recorded. In general, off-the-shelf BCI technology seems to be able to meet the demands set by the field for unobtrusiveness, portability, and simplicity but it is possible that multi-DoF control may be not feasible or impractical without extensive use of supportive AI and advances in electronics charged with decision-making. Finally, from the investigation's perspective, although our study is one among those featuring adequate sized sample of disabled end-users for BCI technology [72, 73], our approach would benefit from wider subject participation for validation and extraction of statistically significant—and relevant to the advancement of the research field—results.

5. Conclusions

Much of the BCI research endeavor has been shaped around the issue of providing affordable, accurate, and real-time assistive systems for the benefit of the disabled. This piece of research adds onto that envisaged puzzle, but tackling affordability (off-the-shelf component system) of realistic scenarios at a reasonable accuracy. Healthy controls, SCI subjects with positive neurological outcome, and subjects with cervical injuries performed better in BCI control. Depression and motivation could play significant roles in BCI and robotics control. Reported imagery capacity was nonsignificantly lower for SCI subjects. User perception of the robot did not differ between SCI and healthy subjects, but, among SCI participants, those with negative outcome rated anthropomorphic characteristics higher. More BCI

sessions are expected to improve performance of SCI and healthy subjects. The herein results demonstrate that by developing BCI decoding algorithms capable of true control of multiple DoFs and addressing the computational cost of online implementation of such an approach, it could be within short-term reach. While maintaining requirements for wireless, unobtrusive solutions constitutes challenges yet to be met, the exploitability of such a system by real patients at a reasonable accuracy cannot be understated. Improving electronics, fluidity, and accuracy of the robotic system, enhancing HRI experience, and implementing a source-based BCI algorithm for multiple class control, as well as further investigations with disabled users, are among our next steps.

Ethical Approval

This study was conducted in accordance with the Declaration of Helsinki (1964) and its following amendments. The Bioethics & Ethics Committee of Faculty of Medicine, Aristotle University of Thessaloniki, approved the study.

Consent

All experiments were conducted with the subjects' understanding and written informed consent.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported in part by 2013 Mario Boni Grant from Cervical Spine Research Society-European Section, awarded to CSI: Brainwave Project. This work was supported in part by the European Union's Horizon 2020 UNCAP project (Grant no. 643555).

Supplementary Materials

Report of statistical analysis. Table 1: descriptive statistics of age for both groups (healthy, SCI). Figure 1: the education level of the healthy and SCI participants. Figure 2: BMI categories across groups (the percentage of participants in each category is displayed on the bars). Figure 3: SCI subjects' age at injury. Table 2: SCI subjects' weight at injury. Figure 4: reported causes of injury. Figure 5: grouping the reported cause of injury by SCI subjects' age. Figure 6: SCI subjects' injury severity as assessed by ASIA Impairment Scale. Figure 7: neurological level of SCI subjects' injuries. Table 3: neurological evaluation scores of the SCI group. Figure 8: median scores of both groups at g-SCIM-III test and its subcategories. Figure 9: smoking status of both healthy and SCI group. Figure 10: performance of both groups at VVIQ test and its subcategories. Figure 11: scores at VVIQ questionnaire and its subcategories of both SCI outcome subgroups. Figure 12: scores of both groups at Beck Depression Inventory (on left hand) and Rosenberg Self-Esteem Scale. Figure 13: BCI performance of both groups in the control of right hand (on the left) and both hands (on the right). Figure 14: BCI scores in SCI group depending on their ASIA classification. Figure 15: BCI performance of SCI group depending on the level of injury. Figure 16: marginally negative correlation between total BCI scores and NLI. Figure 17: mean training scores of both groups at left and right hand. Figure 18: training scores of SCI groups depending on their ASIA classification. Figure 19: mean BCI and training performance depending on the level of injury. Figure 20: answers provided to Godspeed questionnaire from both groups. Figure 21: scores of SCI subjects at GDSPD-Anthr depending on outcome. Figure 22: BCI performance of both groups in different categories of movements. *(Supplementary Materials)*

References

[1] L. R. Hochberg, D. Bacher, B. Jarosiewicz et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, 2012.

[2] J. L. Collinger, B. Wodlinger, J. E. Downey et al., "High-performance neuroprosthetic control by an individual with tetraplegia," *The Lancet*, vol. 381, no. 9866, pp. 557–564, 2013.

[3] A. R. C. Donati, S. Shokur, E. Morya et al., "Long-term training with a brain-machine interface-based gait protocol induces partial neurological recovery in paraplegic patients," *Scientific Reports*, vol. 6, no. 1, p. 30383, 2016.

[4] M. Alam, W. Rodrigues, B. N. Pham, and N. V. Thakor, "Brain-machine interface facilitated neurorehabilitation via spinal stimulation after spinal cord injury: Recent progress and future perspectives," *Brain Research*, vol. 1646, pp. 25–33, 2016.

[5] R. R. Harrison, R. J. Kier, C. A. Chestek et al., "Wireless neural recording with single low-power integrated circuit," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 4, pp. 322–329, 2009.

[6] A. V. Nurmikko, J. P. Donoghue, L. R. Hochberg et al., "Listening to brain microcircuits for interfacing with external world—progress in wireless implantable microelectronic neuroengineering devices," *Proceedings of the IEEE*, vol. 98, no. 3, pp. 375–388, 2010.

[7] J. L. Collinger, S. Foldes, T. M. Bruns, B. Wodlinger, R. Gaunt, and D. J. Weber, "Neuroprosthetic technology for individuals with spinal cord injury," *The Journal of Spinal Cord Medicine*, vol. 36, no. 4, pp. 258–272, 2013.

[8] P. H. Peckham and J. S. Knutson, "Functional electrical stimulation for neuromuscular applications," *Annual Review of Biomedical Engineering*, vol. 7, pp. 327–360, 2005.

[9] L. Pignolo, "Robotics in neuro-rehabilitation," *Journal of Rehabilitation Medicine*, vol. 41, no. 12, pp. 955–960, 2009.

[10] B. He, Ed., *Modeling, Imaging of Bioelectrical Activity*, Springer US, Boston, MA, USA, 2005.

[11] P. L. Nunez and R. Srinivasan, "Electric Fields of the Brain: The neurophysics of EEG," *Electric Fields of the Brain: The neurophysics of EEG*, pp. 1–611, 2009.

[12] A. Athanasiou, I. Xygonakis, N. Pandria et al., "Towards Rehabilitation Robotics: Off-the-Shelf BCI Control of Anthropomorphic Robotic Arms," *BioMed Research International*, vol. 2017, pp. 1–17, 2017.

[13] S. Baillet, "Magnetoencephalography for brain electrophysiology and imaging," *Nature Neuroscience*, vol. 20, no. 3, pp. 327–339, 2017.

[14] J. Meng, S. Zhang, A. Bekyo, J. Olsoe, B. Baxter, and B. He, "Noninvasive electroencephalogram based control of a robotic arm for reach and grasp tasks," *Scientific Reports*, vol. 6, no. 1, article 38565, 2016.

[15] Y. Kim and A. M. Cook, "Manipulation and Mobility Aids," in *Electronic Devices for Rehabilitation*, J. G. Webster, Ed., Wiley, London, 1985.

[16] M. Hillman, "2 Rehabilitation Robotics from Past to Present A Historical Perspective," in *Advances in Rehabilitation Robotics*, pp. 25–44, Springer Berlin Heidelberg, 2006.

[17] W. Seamone and G. Schmeisser, *Evaluation of the APL/JHU Robot Arm Work Station*, 1986.

[18] C. P. Mason and E. Peizer, "Medical Manipulator for Quadriplegics, Colloques IRIA," in *Proceedings of the in International Conference on Telemanipulators for the Physically Handicapped*, pp. 309–312, 1978.

[19] H. H. Kwee, M. Tramblay, R. Barbier et al., "First experimentation of the spartacus telethesis in a clinical environment," *Paraplegia*, vol. 21, no. 5, pp. 275–286, 1983.

[20] Y. Wakita, W.-K. Yoon, and N. Yamanobe, "User evaluation to apply the robotic arm RAPUDA for an upper-limb disabilities Patient's Daily Life," in *Proceedings of the 2012 IEEE International Conference on Robotics and Biomimetics, ROBIO 2012*, pp. 1482–1487, China, December 2012.

[21] V. Maheu, P. S. Archambault, J. Frappier, and F. Routhier, "Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities," in *Proceedings of the Rehab Week Zurich 2011 - 2011 IEEE International Conference on Rehabilitation Robotics, ICORR 2011*, Switzerland, July 2011.

[22] M. S. Johannes, J. D. Bigelow, J. M. Burck, S. D. Harshbarger, M. V. Kozlowski, and T. Van Doren, "An overview of the developmental process for the modular prosthetic limb," *Johns Hopkins APL Technical Digest*, vol. 30, no. 3, pp. 207–216, 2011.

[23] J. J. Daly and J. R. Wolpaw, "Brain-computer interfaces in neurological rehabilitation," *The Lancet Neurology*, vol. 7, no. 11, pp. 1032–1043, 2008.

[24] M. Gomez-Rodriguez, M. Grosse-Wentrup, J. Hill, A. Gharabaghi, B. Scholkopf, and J. Peters, "Towards brain-robot interfaces in stroke rehabilitation," in *Proceedings of the IEEE International Conference on Rehabilitative Robotics*, vol. 2011, 2011.

[25] K. K. Ang, C. Guan, K. S. Chua et al., "A clinical study of motor imagery-based brain-computer interface for upper limb robotic rehabilitation," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '09)*, pp. 5981–5984, Minneapolis, Minn, USA, September 2009.

[26] K. Dautenhahn, "Methodology & themes of human-robot interaction: a growing research field," *International Journal of Advanced Robotic Systems*, vol. 4, no. 1, pp. 103–108, 2007.

[27] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 71–81, 2009.

[28] O. Christ and M. Reiner, "Perspectives and possible applications of the rubber hand and virtual hand illusion in non-invasive rehabilitation: Technological improvements and their consequences," *Neuroscience & Biobehavioral Reviews*, vol. 44, pp. 33–44, 2014.

[29] G. Arfaras, A. Athanasiou, P. Niki et al., "Visual Versus Kinesthetic Motor Imagery for BCI Control of Robotic Arms (Mercury 2.0)," in *Proceedings of the 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 440–445, Thessaloniki, June 2017.

[30] A. Athanasiou, G. Arfaras, I. Xygonakis et al., "Commercial BCI Control and Functional Brain Networks in Spinal Cord Injury: A Proof-of-Concept," in *Proceedings of the 2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 262–267, Thessaloniki, June 2017.

[31] "Brainwave control of a wearable robotic arm for rehabilitation and neurophysiological study in cervical spine injury (CSI:Brainwave)," https://clinicaltrials.gov/ct2/show/NCT02443558.

[32] N. Moustakas, P. Kartsidis, A. Athanasiou, A. Astaras, and P. D. Bamidis, "Development of MERCURY version 2.0 robotic arms for rehabilitation applications," in *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '15)*, July 2015.

[33] S. C. Kirshblum, S. P. Burns, F. Biering-Sorensen et al., "International standards for neurological classification of spinal cord injury (revised 2011)," *The Journal of Spinal Cord Medicine*, vol. 34, no. 6, pp. 535–546, 2011.

[34] R. W. Bohannon and M. B. Smith, "Interrater reliability of a modified Ashworth scale of muscle spasticity," *Physical Therapy in Sport*, vol. 67, no. 2, pp. 206-207, 1987.

[35] J. T. C. Hsieh, D. L. Wolfe, W. C. Miller, and A. Curt, "Spasticity outcome measures in spinal cord injury: Psychometric properties and clinical utility," *Spinal Cord*, vol. 46, no. 2, pp. 86–95, 2008.

[36] M. Itzkovich, I. Gelernter, and F. Biering-Sorensen, "The Spinal Cord Independence Measure (SCIM) version III: Reliability and validity in a multi-center international study," *Disability and Rehabilitation*, vol. 29, no. 24, pp. 1926–1933, 2007.

[37] A. Athanasiou, A. Alexandrou, E. Paraskevopoulos, N. Foroglou, A. Prassas, and P. D. Bamidis, "Towards a Greek adaptation of the Spinal Cord Independence Measure (SCIM)," in *in proceedings of the 15th European Congress of Neurosurgery (EANS 14)*, pp. 181–184, 2015.

[38] D. F. Marks, "Directions in Mental-Imagery Research," *Journal of Mental Imagery*, vol. 19, no. 3-4, pp. 153–167, 1995.

[39] A. T. Beck, R. A. Steer, and M. G. Garbin, "Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation," *Clinical Psychology Review*, vol. 8, no. 1, pp. 77–100, 1988.

[40] W. Petersen, "Society and the Adolescent Self-Image. Morris Rosenberg. Princeton University Press, Princeton, N.J., 1965. xii + 326 pp. $6.50," *Science*, vol. 148, no. 3671, p. 804, 1965.

[41] M. Giannakou, P. Roussi, M.-E. Kosmides, G. Kiosseoglou, A. Adamopoulou, and G. Garyfallos, "Adaptation of the beck depression inventory-II to greek population," *Hellenic Journal of Psychology*, vol. 10, no. 2, pp. 120–146, 2013.

[42] C. Galanou, M. Galanakis, E. Alexopoulos, and C. Darviri, "Rosenberg Self-Esteem Scale Greek Validation on Student Sample," *Psychology*, vol. 05, no. 08, pp. 819–827, 2014.

[43] A. Astaras, A. Athanasiou, A. Alexandrou, P. Kartsidis, N. Moustakas, and P. D. Bamidis, "Double-blind greek translation

and online implementation of the Godspeed robotics questionnaire," in *Proceedings of the in 6th Panhellenic Conference on Biomedical Technology Conference*, p. 34, 2015.

[44] E. I. Konstantinidis, A. Billis, C. Bratsas, A. Siountas, and P. D. Bamidis, "Thessaloniki active and healthy ageing living lab: The roadmap from a specific project to a living lab towards openness," in *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '2016)*, pp. 1–4, July 2016.

[45] E. I. Konstantinidis and P. D. Bamidis, "Density based clustering on indoor kinect location tracking: A new way to exploit active and healthy aging living lab datasets," in *Proceedings of the 2015 IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 1–6, November 2015.

[46] E. I. Konstantinidis, P. E. Antoniou, G. Bamparopoulos, and P. D. Bamidis, "A lightweight framework for transparent cross platform communication of controller data in ambient assisted living environments," *Information Sciences*, vol. 300, no. 1, pp. 124–139, 2015.

[47] T. Sollfrank, D. Hart, R. Goodsell, J. Foster, and T. Tan, "3D visualization of movements can amplify motor cortex activation during subsequent motor imagery," *Frontiers in Human Neuroscience*, vol. 9, article 463, 2015.

[48] D. Cramer, "Fundamental statistics for social research," in *Step-by-step calculations and computer techniques using SPSS for Windows*, Routledge, London, UK, 1988.

[49] D. Cramer and D. Howitt, *The SAGE Dictionary of Statistics*, SAGE Publications, 2004.

[50] D. P. Doane and L. E. Seward, "Measuring skewness: a forgotten statistic?" *Journal of Statistics Education*, vol. 19, no. 2, pp. 1–18, 2011.

[51] N. M. Razali, Y. B. Wah, and M. Sciences, "Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov," *Lilliefors and Anderson-Darling tests*, vol. 2, no. 1, pp. 21–33, 2011.

[52] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality: Complete samples," *Biometrika*, vol. 52, pp. 591–611, 1965.

[53] N. Mogey, "So You Want to Use a Likert Scale?" in *Edinburgh: 2 Evaluation Cookbook Learning Technology Dissemination Initiative*, Institute for Computer Based Learning, Heriot-Watt University, 1998.

[54] N. Birbaumer and P. Sauseng, "Brain–Computer Interface in Neurorehabilitation," 2010.

[55] A. Vuckovic, J. A. Pineda, K. LaMarca, D. Gupta, and C. Guger, "Interaction of BCI with the underlying neurological conditions in patients: pros and cons," *Frontiers in Neuroengineering*, vol. 7, 2014.

[56] F. Nijboer, N. Birbaumer, and A. Kübler, "The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis - a longitudinal study," *Frontiers in Neuroscience*, vol. 4, article no. 55, 2010.

[57] E. M. Hammer, S. Halder, B. Blankertz et al., "Psychological predictors of SMR-BCI performance," *Biological Psychology*, vol. 89, no. 1, pp. 80–86, 2012.

[58] C. Papadelis, C. Kourtidou-Papadeli, P. Bamidis, and M. Albani, "Effects of imagery training on cognitive performance and use of physiological measures as an assessment tool of mental effort," *Brain and Cognition*, vol. 64, no. 1, pp. 74–85, 2007.

[59] F. González-Palau, M. Franco, P. D. Bamidis et al., "The effects of a computer-based cognitive and physical training program in a healthy and mildly cognitive impaired aging sample," *Aging & Mental Health*, vol. 18, no. 7, pp. 838–846, 2014.

[60] S. M. Roldan, "Object Recognition in Mental Representations: Directions for Exploring Diagnostic Features through Visual Mental Imagery," *Frontiers in Psychology*, vol. 8, 2017.

[61] E. Broadbent, "Interactions with Robots: The Truths We Reveal about Ourselves," *Annual Review of Psychology*, vol. 68, pp. 627–652, 2017.

[62] E. López-Larraz, F. Trincado-Alonso, V. Rajasekaran et al., "Control of an ambulatory exoskeleton with a brain-machine interface for spinal cord injury gait rehabilitation," *Frontiers in Neuroscience*, vol. 10, 2016.

[63] G. E. Francisco, N. Yozbatiran, J. Berliner et al., "Robot-Assisted Training of Arm and Hand Movement Shows Functional Improvements for Incomplete Cervical Spinal Cord Injury," *American Journal of Physical Medicine & Rehabilitation*, vol. 96, pp. S171–S177, 2017.

[64] A. Jackson and J. B. Zimmermann, "Neural interfaces for the brain and spinal cord - Restoring motor function," *Nature Reviews Neurology*, vol. 8, no. 12, pp. 690–699, 2012.

[65] R. Grech, T. Cassar, J. Muscat et al., "Review on solving the inverse problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 5, article 25, 2008.

[66] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001.

[67] C. J. Holmes, R. Hoge, L. Collins, R. Woods, A. W. Toga, and A. C. Evans, "Enhancement of MR images using registration for signal averaging," *Journal of Computer Assisted Tomography*, vol. 22, no. 2, pp. 324–333, 1998.

[68] B. J. Edelman, B. Baxter, and B. He, "EEG source imaging enhances the decoding of complex right-hand motor imagery tasks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 4–14, 2016.

[69] H. Yuan, T. Liu, R. Szarkowski, C. Rios, J. Ashe, and B. He, "Negative covariation between task-related responses in alpha/beta-band activity and BOLD in human sensorimotor cortex: an EEG and fMRI study of motor imagery and movements," *NeuroImage*, vol. 49, no. 3, pp. 2596–2606, 2010.

[70] B. Kamousi, A. N. Amini, and B. He, "Classification of motor imagery by means of cortical current density estimation and Von Neumann entropy," *Journal of Neural Engineering*, vol. 4, no. 2, article no. 002, pp. 17–25, 2007.

[71] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.

[72] G. R. Müller-Putz, I. Daly, and V. Kaiser, "Motor imagery-induced EEG patterns in individuals with spinal cord injury and their impact on brain-computer interface accuracy," *Journal of Neural Engineering*, vol. 11, no. 3, Article ID 035011, 2014.

[73] G. Onose, C. Grozea, A. Anghelescu et al., "On the feasibility of using motor imagery EEG-based brain-computer interface in chronic tetraplegics for assistive robotic arm control: A clinical test and long-term post-trial follow-up," *Spinal Cord*, vol. 50, no. 8, pp. 599–608, 2012.

Hindawi
Wireless Communications and Mobile Computing
Volume 2017, Article ID 5062371, 15 pages
https://doi.org/10.1155/2017/5062371

Research Article

Using Smartphones to Assist People with Down Syndrome in Their Labour Training and Integration: A Case Study

Javier Gomez, Juan Carlos Torrado, and Germán Montoro

Department of Computer Engineering, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, Spain

Correspondence should be addressed to Javier Gomez; jg.escribano@uam.es

Received 25 August 2017; Accepted 4 December 2017; Published 21 December 2017

Academic Editor: Evdokimos I. Konstantinidis

This article describes a proposal and case study based on mobile phones and QR Codes to assist individuals with cognitive disabilities in their labour training and integration. This proposal, named AssisT-Task, is a full functional mobile application for Android smartphones and offers step-by-step guidance, establishing a learning method through task sequencing. It has been tested with a group of 10 users and 2 types of labour tasks. Through 7 recorded sessions, we compared the performance and the learning progress with the tool against the traditional assisting method, based on paper instructions. The results show that people with cognitive disabilities learnt and performed better and faster when using AssisT-Task than the traditional method, particularly on tasks that require cognitive effort rather than manual skills. This learning has proved to be essential to obtain an adequate degree of personal autonomy for people with cognitive impairment.

1. Introduction

Autonomy is the main goal for people with any type of cognitive impairment. In order to achieve a fulfilling life, support is critical [1]. In this article, we focus on people with cognitive disabilities susceptible to be recruited to work in a company and, particularly, people with Down syndrome. These people are usually educated in labour centres, where special education teachers and labour tutors train them in several skills, adapted to their profiles. Their curricula usually include internships in companies, and eventually, they are hired afterwards. Therefore, educators have to teach their students a large number of skills, including stationery, office managing, and reception or cleaning service.

The traditional way to train and teach these skills is task repetition for a long time. During this process, caregivers monitor users' performance and provide oral or textual information. Although they usually provide direct supervision and support, they have to let the student progressively complete the task without any help. This way, instruction-based support is usually provided, so that they can consult it and complete the task [2]. Traditional methods of support include verbal instruction, cards with text, and/or pictures or lists. Although this support is carefully developed by caregivers and experts, they present some challenges to people with cognitive disabilities: they often have difficulties to read, to relocate themselves within a text when they get lost, to look for specific information of a certain instruction, and even to understand them.

These issues motivated researchers to study how to introduce technological aids for daily-life activities performance, both in the learning process and in their houses or workplaces. In the literature, we can find that most of the researchers proposed new devices that provide instructions and offer prompting-based interfaces. This approach involves that users must learn how to use a new device—usually wearable screens or specifically developed interfaces—and then use them to do the tasks. In other words, this training involves a new challenge, using an unknown device, which entails a new learning process and changes in their learning paradigm, and its success might be jeopardized again.

Studies such as [3, 4] run different interviews with people with cognitive disabilities and tried to find a common pattern of technology usage among them. Additionally, Hallgrenn et al. [5] applied the ETUQ (Everyday Technology Use Questionnaire) to 120 users with different level of cognitive impairment. It showed that the perceived difficulty of technology use among these people is directly proportional to their cognitive

impairment severity, with a slight influence of the interest of the user in the device or the topic it covers. Lancioni et al. [6] considered that progress on assistive technologies should not be taken apart from progress on common technologies and defended new strategies of intervention so the users get the most of innovative technologies.

Therefore, it would be far better for them to use their well-known devices, in a less intrusive manner, instead of introducing in their lives new and unknown devices. As it will be presented in latter sections, smartphones are very popular among them and fit perfectly these requirements. Even if they had not any previous experience with them before, learning how to use them would be advantageous, since these devices will become useful in some moment of their lives. Besides, Holzinger et al. [7] discussed the acceptance of technology and the tolerance of individuals to introduce a new device in their routine. A person with cognitive disability, yet having been trained in its use, will not use the device in his/her real life if there is no total acceptance of it. The authors concluded that acceptance is strictly related to previous exposure to technology, so smartphones remain as an optimal choice for our study. In fact, choosing a device of high acceptance level decreases the risk to become an abandoned device through time, a problem that occurs with almost 35% of assistive developments, as some studies pointed out [8].

On that basis, we chose smartphones as the developing platform. Our approach was based on an intuitive and uncomplicated way to assist doing tasks. Users only need their smartphones, launch the application, and follow the previously prepared sequence of instructions. Thanks to the ubiquity of these devices, the assistance would be available anytime they need it. The process can be divided into two steps: task selection and task execution. As it will be detailed in further sections, task identification can be challenging so different approaches should be considered. In our case, we decided to release users from this work, tagging the environment with identification codes, easily readable by the device. The cheapest, most widespread, error-safe tagging technology nowadays is QR Codes [9]: task information is coded into a visual mark, which is printed and put near the places where tasks must be done, providing pervasive assistance [10]. Current smartphones are able to read them through the camera and the capturing-decoding procedure is easily includible in applications. On the other hand, the assistance to perform the task will be offered as a prompting sequence of instructions, supplemented with visual and audio cues. This way, users would receive the stimulus by different channels. Besides, thanks to the navigation controls, users would be able to go forward and backward as they need.

In order to validate our development, we run an evaluation with 10 young adults with Down syndrome from a labour training course. We used a hybrid methodology, combining elements from inquiry (e.g., Direct observation during the trials) and test methodologies (e.g., Focus groups), which provided us with objective information about users' performance and knowledge acquisition.

This article is organized as follows: after the introduction, we present a review of the related work in the literature in Section 2. Then, in Section 3, the AssisT-Task system is described in detail. After that, the evaluation carried out is explained in Section 4 and the results are discussed in Section 5. Finally, we summarize the conclusions extracted from the experience and outline some future work lines in Section 6.

2. Related Work

Cognitive disabilities are related to mental and intellectual functioning and can be caused by several factors, such as genetics, congenital, and environmental factors. A wide accepted definition is the one by the American Association on Intellectual and Developmental Disabilities (AAIDD) [11]: "intellectual disability is a disability characterized by significant limitations in both intellectual functioning and in adaptive behaviour, which covers many everyday social and practical skills. This disability originates before the age of 18." Intellectual functioning refers to mental capacity (reasoning, learning, and problem solving), while adaptive behaviour refers to conceptual, social, and practical skills (such as activities of the daily living, occupational skills, and schedules).

Therefore, cognitive disabilities include different diagnosis, such as Alzheimer disease, traumatic brain injury, Down syndrome, autism spectrum disorders (ASD), and attention deficit hyperactivity disorder (ADHD). In this article, we will focus on people with Down syndrome; however, some of the ideas and studies are easily adapted to other cognitive disabilities.

Most of the works related to assistive technologies for people with disabilities are rarely focused on cognitive impairments. Despite this hurdle, the literature still provides some interesting works. Particularly, in this review, we focused on assistive technologies to help people with cognitive disabilities to do different tasks. Due to the wide variety of activities and contexts, we classified the works in three groups: daily-life activities at home, related to education, and related to workplace.

2.1. Daily-Life Activities at Home. As daily-life activities at home, we usually consider every task or basic capability related to personal care (hygiene, dressing, food, etc.), instrumental activities (cleaning, meals preparation, transport, and money management), and the relations with relatives, neighbours, and flat/residence mates. This context provides a basic level of independency and it is narrowly related to age. Therefore, many researches have been working on the empowerment of users at this level. Particularly, a good example of project related to our work is "Memory Aiding Prompting System" (MAPS) [12]. It uses mobile technologies (i.e., PDAs) as assistive devices to present an interactive guide to do activities at home. The guides could be composed of pictures, images, and videos that caregivers prepared with a PC tool. Additionally, the system was extended to MAPS-LifeLine in [13] to provide assistance in the workplace and also feedback to caregivers.

Another interesting project is ARCHIPEL [14], an intelligent kitchen equipped with sensors and actuators developed to help people with specials needs in their daily lives. Thanks to visual and audio cues, users are guided to prepare meals

and alerted in case of dangerous situations (e.g., missing a stove on). Additionally, the environment has a touch screen to guide users doing the activity by providing recipes, videos, and help. Related to this, "The TEeth BRushing Assistance" (TEBRA) project also employs intelligent systems to help users brushing their teeth [10]. It uses sensors, cameras, and complex decision systems to recognize the action that the user performs and automatically provide the next instruction.

Finally, PREVIRNEC is a distributed telerehabilitation system based on virtual environments that allow caregivers to design and adapt activities and rehabilitation programs to users' needs [15]. This process can be done manually by the caregiver or automatically, thanks to the reports generated and the subsequent analysis.

2.2. Related to Education. Another key area that researchers in assistive technologies usually focus on is education. It includes the whole learning process in all stages, from kindergarten to higher education, as well as the relations with other students, teachers, and centres' staff. Besides, innovative technologies are commonly integrated in their curricula as part of the learning areas.

Thus, we found projects such as Artifact-AR [16], an augmented reality system for cognitive rehabilitation. It uses a three-dimensional structure with tags and pointers to work on memory, sequencing, and images identification. The system was evaluated by therapists, who reported its suitability to the rehabilitation and learning processes but also highlighted some limitations of the system, such as its performance-dependency to external illumination and the need of pointers with adapted fasteners.

Related to the use of mobile devices, we found Picaa [17], a mobile platform specifically designed to help people with special needs in their educational process. It takes advantage of the iOS features, such as the touch screen, accessibility, and the ubiquitous access to the Internet to offer four types of activities: visual exploration of content, association, puzzles, and sorting activities. One of the features of Picaa is including the authoring tool and users' activities in the same application, so educators can develop the content directly on the user's device. From the evaluation, the authors concluded that Picaa helped in the development of basic skills, such as perception, attention, or memory.

Another interesting work that uses iOS devices (particularly, iPads) was presented in [18]. In this article, the authors contribute an empirical vision of the iPad as a mobile learning tool for students with cognitive disabilities in postsecondary education. To do that, they distributed devices, previously loaded with applications, among a group of students, and asked about their usage after a period of time. From the surveys and interviews, the authors concluded that, although all users were satisfied, they found difficulties to choose, configure, and test the most suitable application. However, once they found the proper one they reported to be very satisfied.

2.3. Related to the Workplace. One of the problems that people with cognitive disabilities have to address is the lack of autonomy. Authors such as Taylor and Hodapp [1] stated that their real independency relies directly on economic factors. Therefore, finding and keeping a job is a key factor to promote their autonomy. Thus, the workplace context includes all the necessary skills to get, keep, and develop a remunerated work, as well as the personal relations with partners, supervisors, or suppliers.

Thanks to the possibilities that current technologies offer, we can find novel examples of assistive developments such as ARCoach [19], which uses augmented reality to train people with cognitive disabilities in work tasks. Particularly, the authors presented an evaluation of the system applied to meal preparation in a restaurant. To do that, participants had to prepare a menu by selecting four different plates, represented by tags, and putting them in a tray in the appropriate order. A cenital camera recognized the tags and presented a virtual model of each plate on a screen. Then, the menu was automatically analysed by the system, attending to the selection of the plates and their position in the tray. From the evaluation of the system, the authors validated it since participants learnt the activity and this knowledge was retained.

Another example of labour task training based on technology is Kinempt [2]. This system used the Microsoft Kinect camera to identify the action the user performed and compared it to the one that had to be done to complete the task (previously recorded). Once the action is performed correctly, the system automatically provides the next instruction until the user finishes the task. Therefore, no direct supervision is required. The system was validated with real users with cognitive disabilities in a pizzeria environment. They were asked to prepare pizzas with different ingredients with the Kinempt support and the results demonstrated the possibilities and suitability of this kind of developments to make the training period easier and cheaper.

Finally, the literature also provides novel examples of mobile technologies applied to work inclusion, and training. Smith et al. presented in [20] a study of the viability of mobile technologies as self-instruction devices. To do that, three users with cognitive disabilities did a labour related task (upgrading a PC memory) with the support of a mobile phone loaded with videos explaining the procedure. From the data collected, the authors validated these devices as appropriate for labour training. Another interesting work presenting a study of mobile devices (i.e., the iPod touch) as vocational task support was carried out by Gentry et al. [21]. In their article, they present three studies in which participants used an iPod touch loaded with commercial applications as a system of support. During the experiments, the authors measured different time variables, such as the time working, the time they need direct supervision, and external support. In general, thanks to the iPods, the time employed by the labour trainer was drastically reduced as long as the experiment advanced.

In summary, we found that mobile devices (such as smartphones or PDAs) are one of the promising technologies for assistance. Although other technologies such as computer vision or smart environments have been studied, they require an additional and expensive infrastructure. In contrast, the familiarity of users with smartphones, their high penetration level in society, and the increasing capabilities they offer make

them suitable for task assistance in the three contexts: home, education, and workplace. Besides, they act as a motivation, which may help to reduce the high abandonment rates of assistive technologies [8, 22, 23].

In general, the assistance is delivered by different channels or modes: audio, images, video, and so on that can be prompted automatically or by user interaction. That is, some works make the user ask for the next step, while others provide the next instruction automatically. Although both have their advantages and disadvantages, the manual approach would make users more conscious about their progress doing the task, since they have to identify when the step is completed to ask for the next one. This empowers the assistance-training duality [24]. In this sense, the MAPS project fits perfectly with this idea but it has strong limitations with the new devices and interaction possibilities.

Besides, as it has been demonstrated in different studies, mobile devices reduce the supervisor's load, which may lead to costs reductions (both time and human resources).

3. Materials

In the previous section, we have presented a view on people with Down syndrome and current developments for their assistance. As we said before, there were a few approaches that could fit their needs, but they have limitations. Therefore, we decided to design, develop, and evaluate a novel application for Android smartphones.

AssisT-Task is a mobile system based on task-sequencing and QR Codes that provides pervasive guidance to do daily-life activities. To do that, we employ smartphones as prompting devices and performance recorders.

The operation has been simplified as much as possible. First, the caregiver defines the task by means of the set of steps that compose it. This is possible thanks to the authoring tool provided. Once the task definition is ready, a QR Code containing its information is printed and tagged in a proper place (e.g., close to the washing machine for the "doing the laundry" task). Then, users only have to open the application on their phones, point at the tag, and follow the steps to complete the task. Moreover, this guidance is adapted to the task, the user, and his/her needs.

During the activity, the smartphone records every interaction that takes place. This way, caregivers would be able to reproduce and analyse users' performance.

3.1. Data Model and Architecture. The system is based on a client-server architecture. The server stores all the information related to users and activities but the client also has a local copy of the information to provide offline assistance.

Activities (or tasks) are modelled as a set of steps (instructions) and other activities. They also have a name and a unique id, which is coded into a QR so they can be easily identified. On the other hand, each step is represented by a textual instruction and a descriptive image. Besides, they have sorting relations with other steps or tasks to define the sequence. In order to adapt the task to the user and his/her needs, these relations are tagged and two additional features are included into steps: repetitions and branches.

Figure 1: Authoring tool screenshot. The view presents the steps that compose the "prepare coffee" task.

This way, the sequence can be adapted in execution time (the sequence adaptation options are explained in detail in the next subsection).

Caregivers can develop tasks and edit them with a provided authoring tool. It consists of a graphical user interface which allows viewing all the available tasks, modifying them (changing images, descriptions, features as branches, repetitions, or user-labelling), creating new tasks, deleting others, and exporting the QR Code. These features have been implemented into a drag-and-drop environment. An example of the interface is shown in Figure 1. As can be seen, the screen is divided into two parts: a left side bar showing a tree scheme of the available tasks and a right frame presenting the steps that compose the selected task. In this example, the "make coffee" task is composed of seven steps.

Additionally, a toolbar is included in the upper part. It includes drag-and-drop icons to add tasks and steps and buttons to delete and save the work, access user adaptation mode, and generate the task's QR Code.

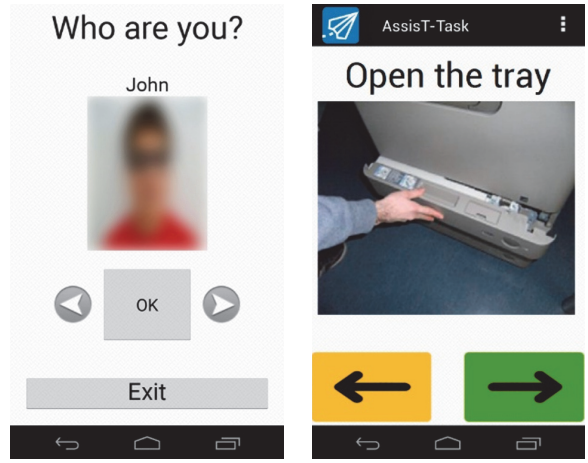
3.2. Sequence Adaptation. Each user has a unique set of abilities, and his/her cognition level is very difficult to measure. Unlike other disabilities, the level of "cognitive prosthesis" given depends on a varying number of factors, where some of them are quite subtle [8]. Therefore, there is an imperative necessity to have flexible tools to assist people with cognitive disabilities. With that idea in mind, AssisT-Task allows caregivers to decrease the prompting level for each user: when designing tasks, each step is chained to the next step by default, but it is possible to label a step specifying what the next step for a certain user is, skipping all steps between them. The intention of this feature is that the designer of the task (usually the caregiver or the job responsible) makes it as granulized as possible, so users with lower cognitive levels have an adequate assistance, whereas users with higher levels can do the same task with less steps, since they do not need so much help. This feature not only covers variability amongst cognitive levels, but also avoids prompting dependency [25] and decreases prompting level a long time. This way, users have an adapted version of the task during their learning process: their caregivers would program to skip steps in their tasks as long as they progress and become able to perform the task with less assistance.

On the other hand, the system also provides mechanisms to adapt the sequence to users' and tasks' contexts. To do that, we developed two mechanisms: repetitions and branches. Under some circumstances, it would be interesting to do a step for a number of times repeatedly. Besides, this number could be specified while designing the task or users should be asked in execution time. This feature is supported by the benefit that people with cognitive disabilities get from mechanical instructions instead of complex or numerical instructions, in most cases. For example, it is preferable to say seal the next envelope repeatedly instead of seal ten envelops for the "prepare the mail" task, for example. The former way makes the step atomic, clear, and understandable by the user, whereas the latter introduces conditional and complex information that can become difficult to understand by some users with a lower cognitive level. In fact, the user will likely assume the complex component of these indications by the time, in a more natural way than a complex instruction.

The other adaptation mechanism is branching. In some daily-life activities, the sequence of steps varies depending on some events during the performance or factors affecting the nature of the activity. For example, doing the laundry is different if clothes are coloured, white, or delicate. With a linear model of steps, caregivers would have to design three different tasks, generate three QR Codes, and put them near the washing machine. With three options, it does not seem very problematic, but with other many activities, photocopying, regarding all the options of paper, zoom, density, and arrangement, it is rather impractical. Therefore, the application allows creating branches in the sequence, through steps that ask the user to choose an option to continue. For example, doing the laundry would have a step asking what kind of clothes you want to wash, with three possible answers: white, coloured, and delicate. Every option would lead to different subsequences of steps, converging before the end, not necessarily, depending on the activity we want to design, with the steps to turn off the washing machine.

3.3. Interface Design. The interface design process was an expert-centred approach. That is, we had the support of experts and therapists of the Down Syndrome Foundation of Madrid and discussed different versions of the interface. It finally took three iterations until we designed the last version of the interface. Figure 2 shows two screenshots of the interface. In (a), the user selection screen is presented. In the first steps of the design, we did not consider the smartphone to be shared by different users but educators suggested that it would be very useful in the class. This way, all users would have the opportunity to use it. Therefore, we decided to include this optional screen to ask the user about who he/she is in order to identify him/her and provide a personalized assistance and registry. As can be seen, the interface has been simplified as much as possible. On the top of the screen, the interface presents users' names and their photographs. Then, there are two arrows to go to the next/previous user and an OK button to get in the application. Additionally, an exit button has been included at the bottom of the screen.

Once the user selects himself/herself, the QR Code screen is launched. It loads a view of the camera and automatically



(a) (b)

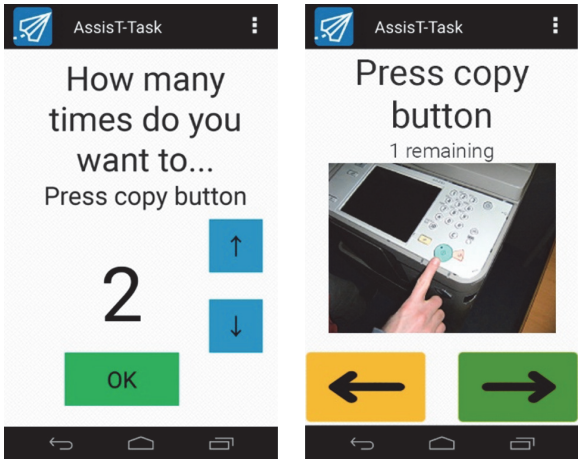
Figure 2: Screenshots of users' selection interface (a). Note that user's photograph has been blurred in order to preserve his privacy. And "open the tray" step (b).

detects and decodes the QR it is pointing at. After that, the system requests the related information and loads the sequence of steps. An example of it is shown in Figure 2(b). On the top of the screen, right under the black title bar, the instruction is shown. The font selected is clean and big enough to be read easily. Right under the text, the descriptive image is included, taking most of the available space. On the bottom of the interface, we included two navigational buttons, to go to the next or the previous step. They have distinct colours, colour-blind-proof, and with a subtle intention: the next button is green, as a metaphor of positive reinforcement, since pressing means advancing within the whole task, whereas the previous button is yellow, meaning a neutral connotation; it is not negative to go back and retry if you feel lost but you have not advanced in the process. The meaning of both buttons is given by the arrow symbols on them. Educators explained that users tend to respond well to arrow indications representing directional messages. Finally, in addition to the textual instruction and the image, the interface can be configured to automatically read aloud the instruction on load.

As it has been said, all the design process and elements of the interface have been carefully studied and discussed with experts in special education. Thus, we have the following.

Texts: they must be shown in a simple and natural style, as recommendation for the caregiver in charge of modelling the task, so it does not become a challenge to the user to understand it. Studies such as [26, 27] showed that, in a moderate level of cognitive disability, text-based instructions become more useful than others that require interpretation of the information or metaphor understanding, if given properly. In fact, nontechnology support for daily-life activities is often given in text-only instruction format. Finally, reading also increases focusing and attention.

Pictures: unless they are clear enough by themselves, they should be highlighted at the zones that the user must pay



(a) (b)

Figure 3: Repetitions screens. (a) Shows the number of repetitions selection interface. (b) Presents the screen of the step that has to be repeated.

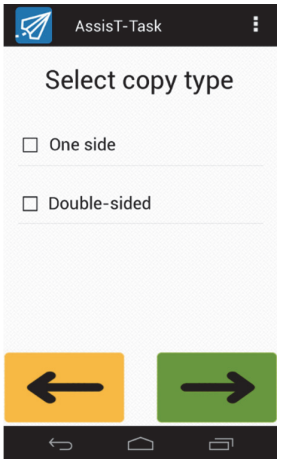


Figure 4: User interface of a branch step. It includes the instruction (select copy type) and the list of options, "one side" and "double sided."

special attention [28]; for example, if the step is turning on the copier, the attached picture would be the control panel of the copier, with a highlighted area around the on/off button.

Audio: when users reach a certain step, the text shown with the description is also read by a text-to-speech engine. It is also read when they touch the screen and when a certain amount of time passes and they have not interacted with the application. Spoken instructions are proved to be the most helpful prompting source in several study cases [27, 29].

Vibration: the device vibrates slightly when a configurable timeout expires. Although some studies tried to build prompting systems only based on vibration of the device, Mechling et al. [27] showed that vibration works better as a supplement of prompting systems based on multimodal support, so here it is used only to notify the user that there has passed some time.

There are two particular cases related to steps' interfaces: repetitions and branches. As it was said before, the number of repetitions can be asked in execution time or set in design time. In the first case, we first present a new screen asking for the number of times to repeat the step. An example of this interface is shown in Figure 3(a). As can be seen, the system asks the user about the number of times she has to do the step: in this example, "press copy button." With the blue buttons with up and down arrows they can choose the appropriate number and then press OK to continue. After that, the step screen is presented. It is pretty similar to the standard one, but a new message has been included, indicating the number of remaining repetitions. The corresponding example is shown in Figure 3(b). As can be seen, under the instruction the system indicates the number of additional repetitions, in a smaller and lighter font.

On the other hand, branches are implemented as lists. The interface shows the instruction as any other step and, instead of the image, we included a list with the options the caregiver designed. An example of an interface is shown in Figure 4. On the top of the screen, there is the instruction: in this case "select copy type." In general, it is advisable to write it as a selection order. Under it, the list of options is shown. In this case, as part as the "make photocopies" task, users would have to choose between one or double-sided copies. Depending on their decision, the sequence would be different.

3.4. Interaction. The interaction with the application has been designed to be the simplest and the most intuitive possible. Apart from the workload required to understand the information that represents a step, users only have to navigate to the next or previous step, in a natural way. The buttons' layout has been designed to be handy and comfortable while holding the mobile phone with both hands or one hand; it is a desirable option for users that had acquired good handling level with the device and are able to perform a task with one hand, while holding the device with the other.

During the early stages of the design process, we run different trials to test whether our designs were suitable for the users. One of the main issues we observed was that some users pressed the next button repeatedly, even without reading or listening the information, thinking that it would allow them to finish earlier, leading them to misunderstandings and errors in the performance. Therefore, we introduced a short delay before the navigation buttons are enabled: users cannot quickly press the next buttons and finish the task; they have to wait for two seconds (this value can be configurable) to be allowed to go to the next step. This way, they are forced to wait and to pay attention to the information presented on the screen.

People with cognitive disabilities often get stunned when they have to remember how to do a task or when understanding an instruction is a bit more complex. Even more, they usually get blocked and cannot continue; therefore, the caregiver has to intervene and provide some stimuli for the user to continue. Regarding that, the application has been designed following a proactive philosophy, so the device not only expects interaction from the user but also requests it.

If certain time has passed since step information is given, the device vibrates and reads aloud again the information; this behaviour is also configurable. This way, we encourage the user to try again to complete the step or draw his/her attention.

4. Evaluation

Even though the design process was assisted by experts, which reduces the probability of technology rejection, we decided to carry out an evaluation with users with cognitive disabilities. We wanted to evaluate the system from two perspectives: the first one was related to its suitability in users' daily lives and the quality of the assistance provided. On the other hand, we wanted to compare users' performance with our system in contrast to traditional support. In general, these methods include paper or cards with actions and pictures, verbal instructions, and direct supervision. They present advantages but, as we said before, many disadvantages, such as the costs (in terms of human resources) and the difficulties that many users present to find information or to recover from an error.

4.1. Methodology. The methodology carried out can be considered as a hybrid between inquiry and test methods [30]. The former involves interacting and observing users during the experiments, group sessions, surveys, and statistical analysis of record files (typically log registries). The later focuses on retrospective tests, such as video analysis, thinking aloud, coaching method, and analysis of users' actions. In general, inquiry methods involve interacting and observing users during the experiments, which may lead to new design ideas and allows a better initial adjustment of the design and usability issues avoidance. On the other hand, test methods involve analysing the results to obtain general conclusions.

The evaluation process took place in a real working place setting. Specifically, all the sessions took place in a labour training centre. It is furnished as an office, with computers, bookcases, file cabinets, and shelves. Additionally, there are office-related devices, such as photocopiers, bind and lamination machines, and recycling points. This setting improves the training activities in the common tasks users may develop in their working places.

In order to reduce the carry-over effect during the sessions, we prepared an incomplete factorial experiment design [31], based on the Latin-square [32], but including repeated measures; this is, participants repeated the tasks several times, in order to get trained as they usually do with traditional methods of support. To do that, we asked the educators and labour trainers of the centre for two different tasks. They selected make photocopies (including configuration options) and documents archiving (according to different criteria). Regarding users profiling, we asked to recruit participants so they could be divided into two equivalent but heterogeneous groups (A and B) to avoid age, gender, or level biases.

Therefore, following the Latin-square experiment design, we assigned tasks, users, and support alternatively. Table 1 summarizes the distribution of tasks, support, and groups. As can be seen, each group did each task using one support alternatively. This way we ensured that results were not biased by users' distribution or the task assigned.

Table 1: Tasks, support, and groups distribution.

	Task 1	Task 2
Group A	AssisT-Task	Paper support
Group B	Paper support	AssisT-Task

Table 2: Tasks, support, and groups distribution during the experiments' weeks.

	1		2-3		4		5-6		8	
	T1	T2	T1	T2	T1	T2	T1	T2	T1	T2
A	X	X	AT	PS	X	X	AT	PS	X	X
B	X	X	PS	AT	X	X	PS	AT	X	X

Each user did each task once a week, during an eight-week period. We arranged with the centre to program the sessions during their workshop classes. This way it would fit perfectly with their curricula and users would not feel under pressure or persuaded. Therefore, we agreed to program photocopies sessions on Monday mornings and archiving sessions on Wednesday mornings. Each participant did the task individually so there was no interference or interaction between participants.

Although support was previously assigned, we introduced some modifications in order to get reference values: during the sessions of the first, the fourth, and the eighth weeks, users performed the tasks without any support but oral instructions at the beginning. Moreover, we left one week (number 7) without training before the last session to check whether the knowledge was acquired and kept in time or not. Therefore, the distribution of tasks support and groups is summarized in Table 2. Each week (from 1 to 8) is divided into two columns, the first corresponding to Task 1 (photocopies) and the second to Task 2 (documents archiving). Each row corresponds to a group (A or B) and the cells present the support used (X means no support, only oral instructions; PS, paper support; and AT, AssisT-Task). All the sessions were recorded and two observers took notes about users' performance and any other relevant issue.

The number of repetitions came determined by the educators of the centre. After reviewing users' performance on week 6, they decided that most of the users were doing their best so we could conclude the experiment. Despite the limited number of sessions, educators reported that they usually program the same number of sessions but, due to time limitations, they cannot do each task once a week and they have to extend the time between sessions up to once every two weeks for each user. Therefore, we already add a value to our system: it allowed a more efficient way of training regarding time resources.

4.2. Tasks. As it was said before, we carefully chose the tasks by asking experts about the most suitable ones. Thus, we prepared a list of requisites:

(1) These tasks should be interesting for users, both from the user point of view (enjoyable) and from the

formation curricula point of view. This way, users may understand the experiment as part of their studies which may avoid biases.

(2) They should be easily arranged as a sequence of steps and standardized as much as possible. This is, they should be appropriate for all the participants.

(3) Tasks should be different enough to avoid carry-over effects, but relative similar in terms of difficulty or time needed.

(4) As far as possible, tasks should not have been trained before.

According to these requisites and experts' criteria, the first task (T1) consisted of making photocopies and the second one (T2) was archiving documents. Both tasks fulfilled all the requisites but the fourth. Since the formation course takes two years and many of the participants were on their second semester, they had already been trained on photocopies and, some of them, also on sorting documents (but not archiving). Therefore, we decided to increase T1 difficulty by introducing features configuration on the photocopier: users had to make one (and only one) copy of a map of the subway, but it had to be reduced from a DIN-A3 size to a DIN-A4 and the density had to be increased exactly 3 points. This task was modelled as a sequence of 10 steps, including a branch step: users had to decide the type of copy from a list (simple, double sided, enlargement, or reduction). On the other hand, T2 consisted of documents archiving. In this case, educators prepared a set of contracts and put them unsorted on a desktop tray. Users had to put one of them in a proper file, regarding the following criteria: the document date had to be 2012-2013; then, they had to look for the WOP code (an invented array of letters and numbers). Depending on it, they had to choose the proper folder (there were three possibilities). Finally, they had to look for a company name on the first paragraph of the contract and archive the document alphabetically depending on it. This task was modelled as a sequence of 9 steps, including a branch as well (asking for the folder name).

Although both tasks were pretty similar in terms of number of steps and time to be completed, they were intentionally different regarding the skills required: while T1 required manipulative skills to handle the paper, open and close the machine, and so on. T2 required strong cognitive skills since users had to read, look for concrete information within a text, and sort alphabetically.

Support materials for both tasks were also developed by the educators. For the traditional support (paper based), we used the materials they already employed in their courses. Task 1 manual included instructions to do different type of photocopies on the same sheet of paper. They used highlighted fonts to separate the type of photocopies and numbered the steps of each activity. The traditional support for documents archiving was more elaborated and included examples and colours to highlight relevant information (such as one colour for each folder). On the other hand, AssisT-Task support was developed specifically for this evaluation. They were based on the traditional support (same instructions) but we included photographs as well for both tasks.

4.3. Users Profiling. Participants of labour training programs usually have mild to moderate cognitive disabilities. In many cases, they also have other disabilities, such as reduced vision or mild motor impairments. Despite their disabilities, most of them are able to read, understand simple instructions, do basic calculus, and have social manners and politeness. All these skills are usually acquired in previous stages, and now they focus on the abilities and capabilities typical of the workplace.

Thus, we asked educators to recruit participants regarding each one's capabilities and the possible benefit they could get from the experience. In order to get a wide vision of the field and attend to the diversity, we asked them to select users of different levels so we had some heterogeneity, both genders, and typical age range (around 20 years old). Hence, they chose 10 users, 5 males and 5 females, who were 23.8 years old on average (SD = 4.77). The certified degree of disability oscillates between 33% (the minimum required according to current law) and 75%. Although the certificate is necessary to access special education centres and government support, many educators and specialists rate their users according to different abilities regarding four dimensions: cognition, social skills, handling capabilities, and attitude. Each dimension has a set of characteristics and, each one, is usually rated from 1 to 3 (lower values mean lower capabilities). In relation to this study, the most important characteristics are the ones of cognition and handling capabilities dimensions, which are as follows.

Cognition

(i) Attention: the ability to keep concentrated on an object/action/task

(ii) Memory: the ability to hold and manipulate information in the short or long term

(iii) Instructions comprehension: the capability to understand and process simple and/or complex instructions

(iv) Flexibility

Handling Capabilities

(i) Mobility

(ii) Rhythm

(iii) Cleanliness

The distribution of users in groups and their profiles are summarized in Table 3.

As can be seen, most of the job profiles are office assistants, so they are usually trained on the typical tasks they will have to develop in their work. Therefore, tasks fitted perfectly to their curricula and were very interesting for their formation.

In order to get their technological profile and their familiarity with mobile devices we made an interview. From their answers, we concluded that all participants had a mobile phone, but only 6 out of 10 considered it as a smartphone. Besides, tablets were also popular, but not all of them had one. Some of them reported they used a relative's one. In relation to Internet access, all participants reported they had connection

Table 3: Users profiles.

	User	Age	Gender	Cognitive level			Social skills			Handling capabilities			Attitude		% disability	Job profile
				Attention	Memory	Instructions comprehension	Flexibility	Basics	Relation	Decisive	Mobility	Rhythm	cleanliness	Responsibility	Motivation		
A	1	22	M	3	3	3C	3	3	3	3	3	3	3	3	3	55	Complex office assistant
	2	37	F	3	3	2S	1	3	2	1	2	2	3	3	2	68	Simple office assistant
	3	25	F	1	2	1S	1	1	2	2	3	3	2	2	2	75	Office/education assistant
	4	27	M	3	3	3S	2	2	3	3	3	3	1	3	3	65	Complex office assistant
	5	21	M	2	2	2C	3	3	3	3	3	2	3	2	2	33	Off. assist facing public
B	6	21	F	3	3	3S	2	2	2	3	3	2	3	3	3	75	Office assistant
	7	21	M	3	3	3S	3	3	3	2	1	2	2	3	3	65	Storekeeper, mailing
	8	22	M	3	3	3C	1	3	3	3	3	3	3	3	3	53	Any
	9	22	F	2	3	2S	1	3	3	2	3	2	3	3	3	65	Simple office assistant
	10	23	F	2	3	3S	3	3	3	3	3	2	3	3	3	65	Simple office assistant

at home and part of them also in their phones. Related to smartphone usage, the most popular purposes are instant messaging (i.e., WhatsApp), photography, multimedia reproduction, and videogames. Besides 9 out of the 10 participants reported they use their email frequently and social networks (i.e., Facebook). Thus, and despite their disabilities, they have similar technological profiles to people without disabilities of their same age.

5. Results and Discussion

In order to evaluate AssisT-Task in terms of the quality of assistance and suitability (regarding users' performance improvement), we made a retrospective analysis of the records, phone registries, and observers' notes. Traditionally, experts in special education value users' performance regarding two factors: completion and time needed: that is, if they finish the task properly and in less time. In our case, in agreement with the educators, all users had to finish the task. This is, in case they made a mistake that would not allow them to finish the task correctly, the observer helped them to recover from the error. This decision was made in order to act as they usually do in the centre (with the traditional support).

Regarding data collection, we focus on the following measurements during the analysis:

(1) Completion time: measured as the time between the start and the end of the performance. This factor must be taken into account carefully, due to its weak representativeness when comparing between subjects: the fact that one user takes one minute against another one who takes five minutes does not show an actual difference between performance qualities. In general, some users simply take more time to complete a task than others, regardless of their success in the task. However, this measure becomes an invaluable progress indicator when it is used within subjects, in other words, when comparing the time taken to complete a task in the first session with the last one.

(2) Errors: we counted an error when users did not follow the specified instruction. For example, in T1 we considered an error to make two copies instead of one: the instructions ordered specifically one copy. Considering T2, a typical error was filing a document into an incorrect tab of the folder.

(3) Help requests: that is, the number of times users asked the observers for assistance. In some cases, users did not know how to continue, got lost, or hesitated at some point of the task and asked directly. In other cases, they looked for approval or made gestures to indirectly call the observer (on the view of the experts).

These measurements were analysed for each user and session to study their advance individually (within-subjects analysis). The evolution of each measurement along the sessions is represented in Figures 5, 6, and 7 (completion time, errors, and help requests, resp.). These figures are composed of 4 graphs each, one for every group, task, and support combination. Completion time is shown in Figure 5: (a) corresponds to Group A, T1, and AssisT-Task as support; (b) shows time results as well for T1; but in this case, it corresponds to Group B, who used paper support; and (c) and (d) correspond to T2 and the proper support for both groups. As can be seen, most of the users present a decreasing curve, which means they needed less time to do the tasks as sessions progressed. The exception is graph (c), where this decrease is not so clear. We consider that this issue came motivated by the complexity of the task and the fact that they got confused with the paper instructions. As we will see in later analysis, the other measurements are also unclear for this combination (Group A, T2, and paper support).

In relation to errors, we can observe similar behaviours in Figure 6. There is a tendency to decrease the number of errors made. The exception again is the (c) graph, which is blurry and there is no clear trend. In this case, participants who used AssisT-Task did less mistakes than the other group. At this point, we want to remind that the number of errors (or just the presence of errors) can make users not complete the task correctly, which is one of the most relevant factors to consider.

Finally, help requests are presented in Figure 7. As can be seen in the graphs, there were only a few requests during the sessions. Moreover, as we observed directly and later on the videos, most of them were very subtle and looking for approval. Again, and for both tasks, users who were supported by AssisT-Task asked less for help. This could be motivated by the fact that participants did not get lost when using AssisT-Task.

On the other hand, if we analyse the behaviour between subjects, we can observe that there is evidence of the influence of the support on the results. In Table 4, we have summarized the statistical analysis of the measurements for both tasks. As can be seen, for T1 there is no statistical evidence of the influence of the support, while in T2, it exists. The number of errors and help requests is significantly lower for AssisT-Task ($p < 0.05$). As it was said before, time factors are not as representative by themselves as other factors (such as errors). Making mistakes may lead to not completing the task correctly and it is even more important than doing activities fast. However, due to the limited number of users, statistical analysis should be considered carefully and as an illustrative result.

As it was presented in Methodology, in sessions 1, 4, and 7, users did not have any support but oral instructions at the beginning. In Table 5, we have summarized the average values of time and errors measurements for both tasks and groups (see Table 1 for task, support, and group distribution). Help requests were also analysed, but due to the reduced number of values, it has not been included in this part of the study. Regarding the time needed, there is an improvement for both tasks and groups, except for T2 and group A. That is, on average, all users improved except the ones that trained the cognitive task with paper support. Therefore, paper support seemed to be the less appropriate support to train tasks that require higher cognitive load. On the other hand, if we focus on the average number of errors and T2, there is also a slight

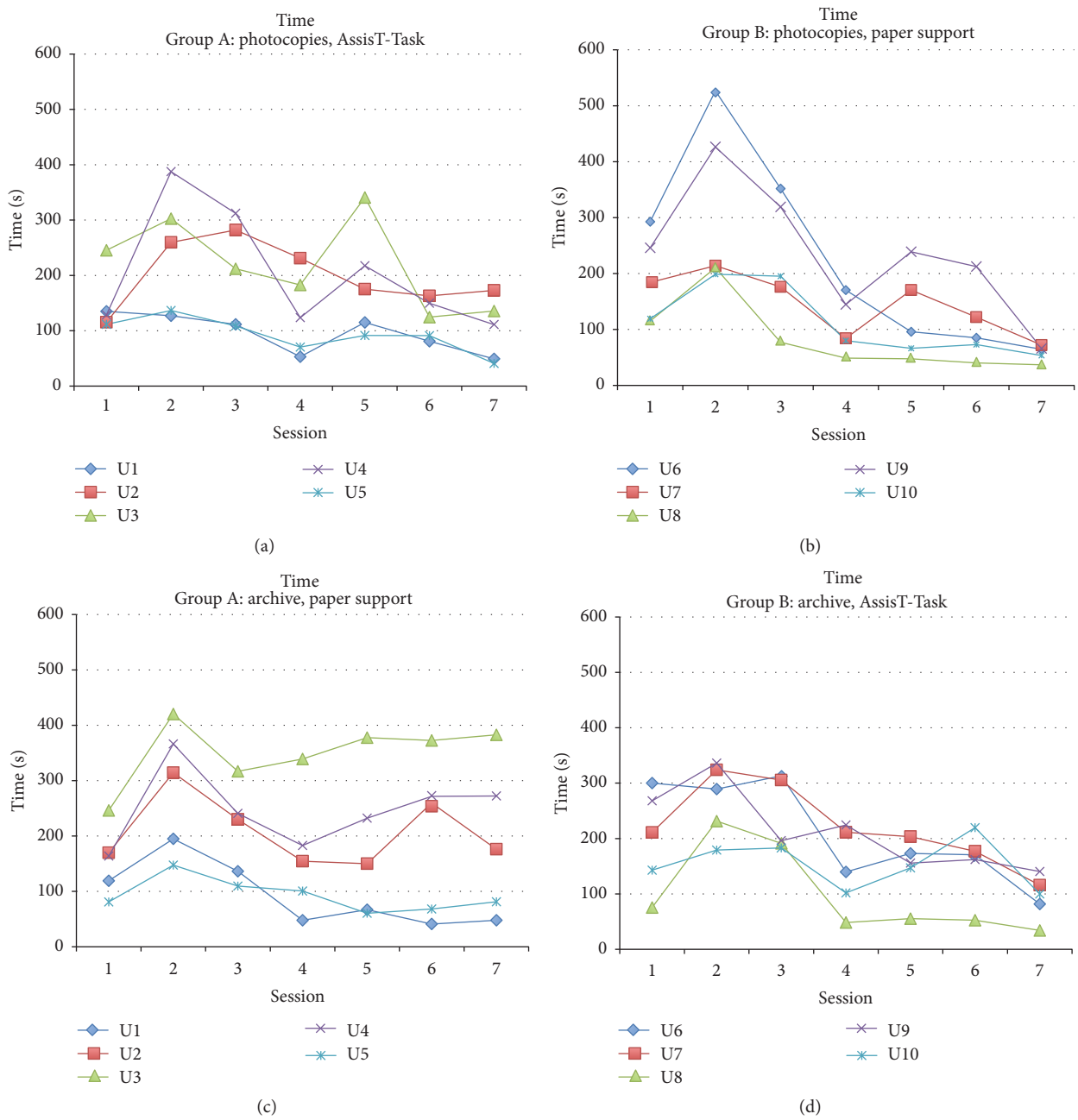


Figure 5: Completion time measurements along the sessions for each user. (a) and (b) present T1 values, while (c) and (d) correspond to T2.

Table 4: Statistical analysis of the data.

	Paper support	AssisT-Task	p value
	Task 1: photocopies		
Time (s)	160.75 (115.34)	162.78 (86.36)	0.463
Errors	0.80 (0.90)	1.11 (1.02)	0.154
Help req.	0.77 (1.46)	0.49 (1.04)	0.248
	Task 2: documents archiving		
Time (s)	198.38 (112.44)	178.87 (83.33)	0.601
Errors	1.8 (1.47)	0.77 (1.06)	**0.001**
Help req.	0.46 (0.74)	0.2 (0.63)	**0.032**

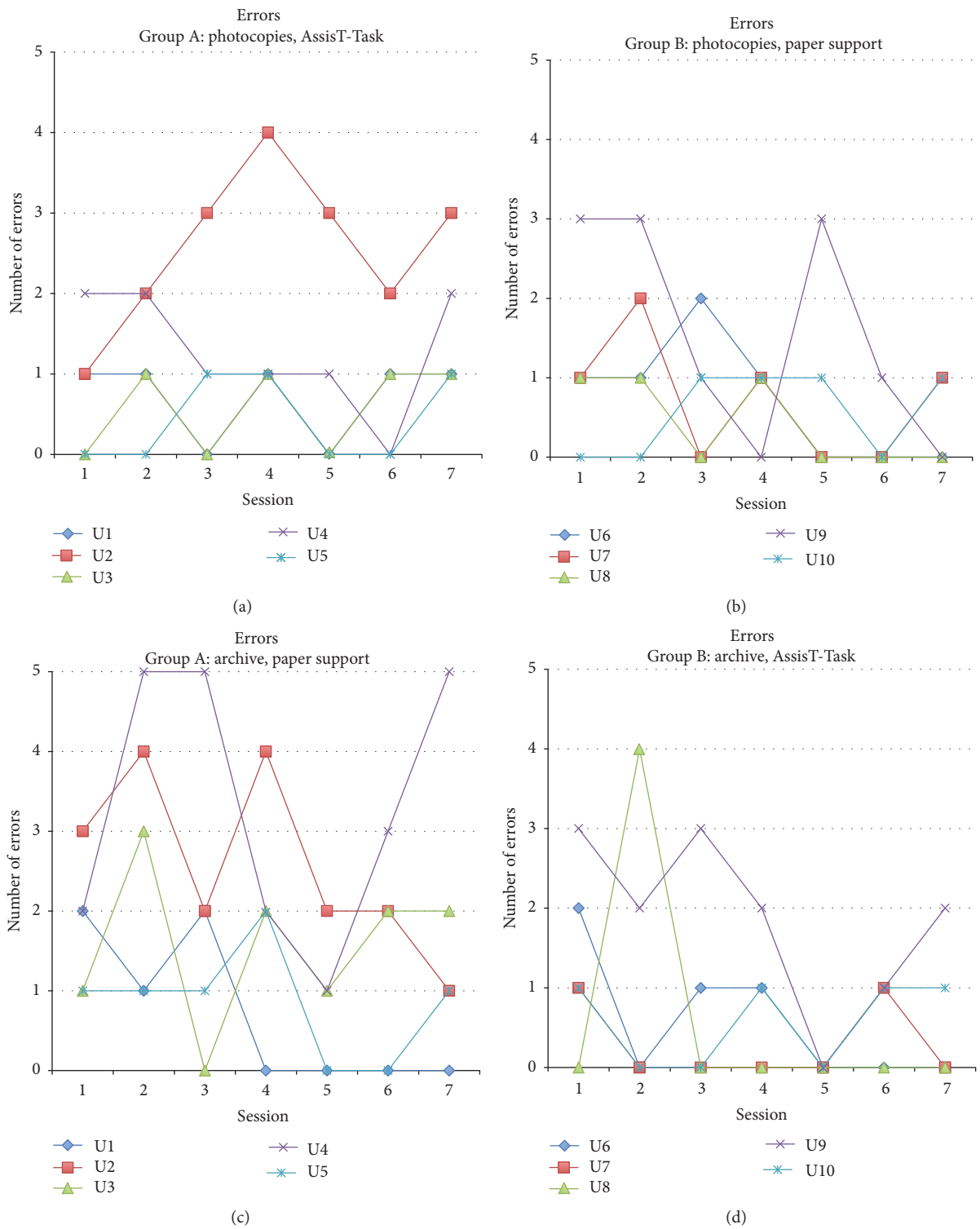


Figure 6: Number of errors along the sessions for each user. (a) and (b) present T1 values, while (c) and (d) correspond to T2.

improvement in Group B, the one that was trained using AssisT-Task.

From the recordings, observers' notes, and focus groups with educators and labour trainers we made a qualitative analysis. In general, all users handled the smartphone properly. That is, all of them hold the smartphone in portrait mode, as the application was designed.

Regarding the QR scanning, all of them understood perfectly the process. Many of them named it as "taking a picture of the code." Therefore, it was easy for them to point

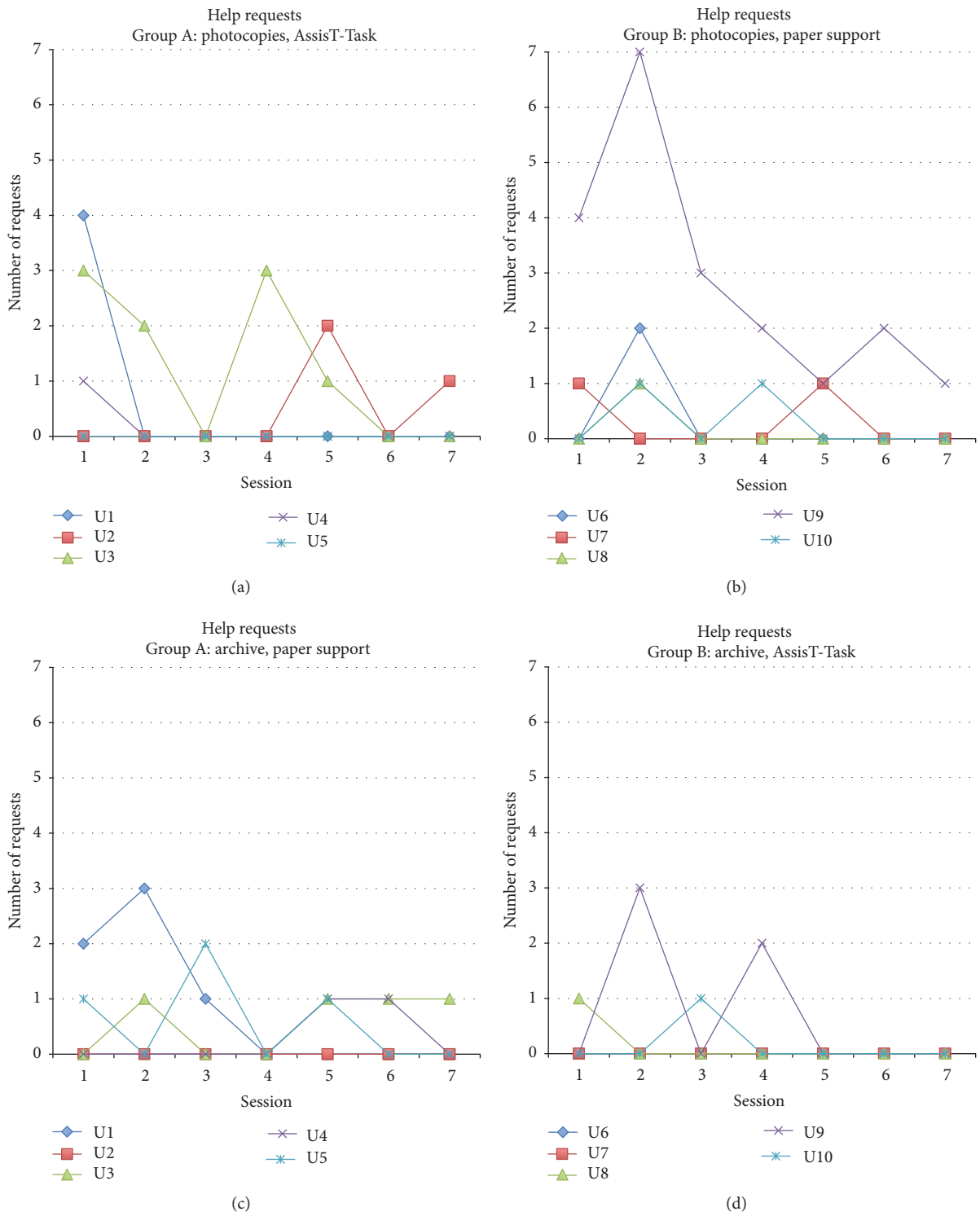


Figure 7: Number of help requests along the sessions for each user. (a) and (b) present T1 values, while (c) and (d) correspond to T2.

at the code with the phone and wait for the application to capture it.

Finally, another interesting conclusion we extracted from the recordings and educators and labour trainers was the motivational component of the application. As young adults, most of them are very interesting in new technologies. Therefore, using them as part of their formation made them keener to participate and do their best.

Table 5: Average values for time and errors measurements of sessions 1, 4, and 7.

	S	Task 1: Photocopies		Task 2: Documents archiving	
		Group A	Group B	Group A	Group B
Time	1	147.20 (55.74)	191.39 (77.91)	155.94 (61.94)	199.545 (91.60)
	4	132.31 (74.98)	105.73 (50.15)	156.05 (101.19)	145.12 (73.96)
	7	101.97 (56.47)	58.62 (13.89)	192.14 (137.89)	94.46 (40.03)
	S	Task 1: Photocopies		Task 2: Documents archiving	
		Group A	Group B	Group A	Group B
Errors	1	0.8 (0.84)	1.2 (1.04)	1.8 (0.84)	1.4 (1.14)
	4	1.6 (1.34)	0.8 (0.45)	2 (1.41)	0.8 (0.84)
	7	1.6 (0.89)	0.4 (0.55)	1.8 (1.92)	0.6 (0.89)

6. Conclusions

Although the evaluation revealed promising results, they should be carefully considered: we tried to include users with different capabilities and levels, which introduces value to the experiment, but due to the limited number of users and variation (in terms of type of disability) and the number of sessions we cannot universalize the study for all people with cognitive disabilities. However, we think it is representative for a particular group: people with Down syndrome who are being trained to get a job.

First of all, we would like to highlight the variety of the results. As can be seen in Table 4, in some measurements the standard deviation is relative high. This denotes the presence of many atypical values and lack of normality. On the other hand, the statistical analysis (presented on the same table) demonstrates the influence of the type of support in relation to the number of errors and help requests in the T2 case. In contrast, we did not find any evidence for T1. This may be motivated by the nature of both tasks: while T1 requires higher manipulative capabilities, T2 requires cognitive skills.

Secondly, and as it was said before, the completion time is usually related to knowledge acquisition although it is not the most representative. We did not find any evidence of the influence of the support in this factor in our study.

In addition to the retrospective analysis, we carried out focus groups with educators and labour trainers. In their opinion, AssisT-Task fitted perfectly for higher and lower profiles. Higher profiles are usually more impulsive and try to finish the tasks quickly, regardless whether they are doing it right or wrong. Moreover, they are reluctant to follow fixed and repetitive orders. This issue can influence their chances to get a job. On the other hand, AssisT-Task was ideal for them. In fact, as it was reported by the educators after the experiment, one of the users (U9) was selected to participate although she had a very low profile and was not valued for cognitive tasks. Surprisingly, she was able to do the archiving task perfectly with the support, and satisfactory without any kind of help. This fact demonstrates that AssisT-Task provides new opportunities for these users.

As future work, we propose to extend the trials, including more users and settings, as well as different tasks. Moreover, it would be very interesting to test the system in a real setting (company) and evaluate the impact of AssisT-Task in the work-inclusion process.

Additionally, educators suggested improving the authoring tool to make it available on tablets. This way all the design process could be done on-site.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors wish to thank the Madrid's Down Foundation staff for their collaboration in this research. The work has been partially funded by the following projects: "e-Training y e-Coaching para la Integración Socio-Laboral" (TIN2013-44586–R) and "eMadrid-CM: Investigación y Desarrollo de Tecnologías Educativas en la Comunidad de Madrid" (S2013/ICE-2715).

References

[1] J. L. Taylor and R. M. Hodapp, "Doing nothing: Adults with disabilities with no daily activities and their siblings," *American Journal on Intellectual and Developmental Disabilities*, vol. 117, no. 1, pp. 67–79, 2012.

[2] Y.-J. Chang, L.-D. Chou, F. T.-Y. Wang, and S.-F. Chen, "A kinect-based vocational task prompting system for individuals with cognitive impairments," *Personal and Ubiquitous Computing*, vol. 17, no. 2, pp. 351–358, 2013.

[3] E. LoPresti, C. Bodine, and C. Lewis, "Assistive technology for cognition," *IEEE Engineering in Medicine and Biology Magazine*, vol. 27, no. 2, pp. 29–39, 2008.

[4] N. M. Gell, D. E. Rosenberg, G. Demiris, A. Z. LaCroix, and K. V. Patel, "Patterns of technology use among older adults with and without disabilities," *The Gerontologist*, vol. 55, no. 3, pp. 412–421, 2015.

[5] M. Hällgren, L. Nygård, and A. Kottorp, "Everyday technology use among people with mental retardation: Relevance, perceived difficulty, and influencing factors," *Scandinavian Journal of Occupational Therapy*, vol. 21, no. 3, pp. 210–218, 2014.

[6] G. E. Lancioni, N. N. Singh, M. F. O'Reilly, J. Sigafoos, and D. Oliva, "Assistive technology for people with severe/profound intellectual and multiple disabilities," in *Assistive Technologies for People with Diverse Abilities*, pp. 277–313, Springer, 2014.

[7] A. Holzinger, G. Searle, and M. Wernbacher, "The effect of previous exposure to technology on acceptance and its importance

in usability and accessibility engineering," *Universal Access in the Information Society*, vol. 10, no. 3, pp. 245–260, 2011.

[8] M. Dawe, "Desperately seeking simplicity: how young adults with cognitive disabilities and their families adopt assistive technologies," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 1143–1152, 2006.

[9] M. Gogova and N. Koceska, "The use of QR codes in education , A journal for information technology, education development and teaching methods of technical and natural sciences," Tech. Rep., University of Novi Sad Technical Faculty, Mihajlo Pupin , Zrenjanin Department of Teaching Methods of Science and Education Technology, 2014.

[10] C. Peters, T. Hermann, S. Wachsmuth, and J. Hoey, "Automatic task assistance for people with cognitive disabilities in brushing teeth-a user study with the tebra system," *ACM Transactions on Accessible Computing (TACCESS)*, vol. 5, no. 4, Article ID 2579700, 2014.

[11] J. R. Thompson, V. J. Bradley, W. H. E. Buntinx et al., "Conceptualizing supports and the support needs of people with intellectual disability," *The American Journal on Intellectual and Developmental Disabilities*, vol. 47, no. 2, pp. 135–146, 2009.

[12] S. Carmien, "Task support for people with cognitive impairments and their caregivers," *American Journal of Occupational Therapy*, vol. 14, no. 3, pp. 1–4, 2004.

[13] S. Carmien, R. Depaula, A. Gorman, and A. Kintsch, "Increasing workplace independence for people with cognitive disabilities by leveraging distributed cognition among caregivers and clients," *Computer Supported Cooperative Work*, vol. 13, no. 5-6, pp. 443–470, 2004.

[14] J. Bauchet, S. Giroux, H. Pigot, D. Lussier-Desrochers, Y. Lachapelle et al., "Pervasive assistance in smart homes for people with intellectual disabilities: a case study on meal preparation," *International Journal of Assistive Robotics and Mechatronics*, vol. 9, no. 4, pp. 42–54, 2008.

[15] D. Tost, S. Grau, M. Ferré et al., "PREVIRNEC: A cognitive telerehabilitation system based on virtual environments," in *Proceedings of the 2009 Virtual Rehabilitation International Conference, VR 2009*, pp. 87–93, Israel, July 2009.

[16] C. Kirner and T. G. Kirner, "Development of an interactive artifact for cognitive rehabilitation based on augmented reality," in *Proceedings of the 2011 International Conference on Virtual Rehabilitation, ICVR 2011*, Switzerland, June 2011.

[17] Á. Fernández-López, M. J. Rodríguez-Fórtiz, M. L. Rodríguez-Almendros, and M. J. Martínez-Segura, "Mobile learning technology based on iOS devices to support students with special education needs," *Computers & Education*, vol. 61, no. 1, pp. 77–90, 2013.

[18] L. Chmiliar and C. Anton, "The iPad as a mobile learning tool for post-secondary students with disabilities," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Preface*, vol. 8548, no. 2, pp. 597–600, 2014.

[19] Y.-J. Chang, Y.-S. Kang, and P.-C. Huang, "An augmented reality (AR)-based vocational task prompting system for people with cognitive impairments," *Research in Developmental Disabilities*, vol. 34, no. 10, pp. 3049–3056, 2013.

[20] K. A. Smith, S. B. Shepley, J. L. Alexander, A. Davis, and K. M. Ayres, "Self-Instruction Using Mobile Technology to Learn Functional Skills," *Research in Autism Spectrum Disorders*, vol. 11, pp. 93–100, 2015.

[21] T. Gentry, S. Lau, A. Molinelli, A. Fallen, and R. Kriner, "The Apple iPod Touch as a vocational support aid for adults with autism: Three case studies," *Journal of Vocational Rehabilitation*, vol. 37, no. 2, pp. 75–85, 2012.

[22] S. P. Carmien, "Socio-technical environments and assistive technology abandonment," 2010.

[23] B. Phillips and H. Zhao, "Predictors of Assistive Technology Abandonment," *Assistive Technology*, vol. 5, no. 1, pp. 36–45, 1993.

[24] A. M. Seelye, M. Schmitter-Edgecombe, B. Das, and D. J. Cook, "Application of cognitive rehabilitation theory to the development of smart prompting technologies," *IEEE Reviews in Biomedical Engineering*, vol. 5, pp. 29–44, 2012.

[25] L. M. Sabielny and H. I. Cannella-Malone, "Comparison of prompting strategies on the acquisition of daily living skills," *Education and Training in Autism and Developmental Disabilities*, vol. 49, no. 1, pp. 145–152, 2014.

[26] M. E. Hudson, D. M. Browder, and B. Jimenez, Effects of a peer-delivered system of least prompts intervention and adapted science read-alouds on listening comprehension for participants with moderate intellectual disability,.

[27] L. C. Mechling, "Assistive technology as a self-management tool for prompting students with intellectual disabilities to initiate and complete daily tasks: a literature review," *Educ. Train. Dev. Disabil*, vol. 42, no. 3, p. 252, 2007.

[28] Y.-J. Chang and T.-Y. Wang, "Comparing picture and video prompting in autonomous indoor wayfinding for individuals with cognitive impairments," *Personal and Ubiquitous Computing*, vol. 14, no. 8, pp. 737–747, 2010.

[29] S. Fickas, M. Sohlberg, and P.-F. Hung, "Route-following assistance for travelers with cognitive impairments: a comparison of four prompt modes," *International Journal of Human-Computer Studies*, vol. 66, no. 12, pp. 876–888, 2008.

[30] J. L. Vidal, J. A. González, and A. I. P. Ordenador, *La Interacción Persona-Ordenador*, AIPO, 2001.

[31] R. E. Kirk, *Experimental Design*, Wiley Online Library, 1982.

[32] G. E. P. Box, J. S. Hunter, and W. G. Hunter, *Statistics for Experimenters: Design, Innovation, and Discovery*, John Wiley, 2005.